Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Nonprogressive Diffusion on Social Networks: Approximation and Applications

Yunduan Lin, Heng Zhang, Renyu Zhang, Zuo-Jun Max Shen

To cite this article:
Yunduan Lin, Heng Zhang, Renyu Zhang, Zuo-Jun Max Shen (2025) Nonprogressive Diffusion on Social Networks: Approximation and Applications. Management Science

Published online in Articles in Advance 01 Dec 2025

. https://doi.org/10.1287/mnsc.2022.03031

https://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–20
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Nonprogressive Diffusion on Social Networks: Approximation and Applications

Yunduan Lin,[a] Heng Zhang,[b] Renyu Zhang,[a,*] Zuo-Jun Max Shen[c,*]

[a] The Chinese University of Hong Kong Business School, The Chinese University of Hong Kong, Hong Kong; [b] W. P. Carey School of Business, Arizona State University, Tempe, Arizona 85287; [c] Faculty of Engineering and Faculty of Business and Economics, University of Hong Kong, Hong Kong

*Corresponding authors

Contact: yunduanlin@cuhk.edu.hk, https://orcid.org/0000-0001-9155-3652 (YL); hengzhang24@asu.edu, https://orcid.org/0000-0002-6105-6994 (HZ); philipzhang@cuhk.edu.hk, https://orcid.org/0000-0003-0284-164X (RZ); maxshen@hku.hk, https://orcid.org/0000-0003-4538-8312 (Z-JMS)

Received: September 28, 2022
Revised: December 5, 2023; March 31, 2025; July 13, 2025
Accepted: July 20, 2025
Published Online in Articles in Advance: December 1, 2025

https://doi.org/10.1287/mnsc.2022.03031

Abstract. Nonprogressive diffusion models the spread of behavior on social networks, where agents are allowed to reverse their decisions as time evolves. To provide an efficient framework for evaluating and optimizing nonprogressive diffusion, we introduce a comprehensive model along with a fixed-point approximation (FPA) scheme, which admits both theoretical guarantee and computational efficiency. We show that the approximation error depends on the network structure and derive order-optimal bounds for this error based on a newly proposed network measure. Additionally, we propose two easy-to-calculate network metrics (one at the node level and the other at the network level) that serve as reliable indicators of FPA performance. Our results indicate that the FPA scheme is particularly accurate for dense and large networks, which are typically challenging to analyze via simulation. To showcase the broad applicability of our approach, we apply the FPA scheme to well-known problems, like influence maximization and optimal pricing on social networks. Finally, we conduct extensive numerical experiments on both synthetic and real-world networks. On real-world networks, the FPA scheme achieves computational speedups of 70–230 times compared with naïve agent-based simulation and 23–30 times compared with a more advanced simulation method while maintaining a mean absolute percentage error of less than 3.48%.

History: Accepted by Jeannette Song, operations management.
Funding: R. Zhang is grateful for financial support from the National Natural Science Foundation of China [Grant 72422004] and the Hong Kong Research Grants Council General Research Fund [Grants 14502722, 14504123, and 14503224].
Supplemental Material: The online appendix and data files are available at https://doi.org/10.1287/mnsc.2022.03031.

Keywords: nonprogressive network diffusion • large-scale network approximation • network centrality • influence maximization • pricing

1. Introduction

Social networks deeply shape our lives. People are more receptive to information from friends (Lu et al. 2013) and more inclined to make purchases based on social referrals (Bapna and Umyarov 2015, Ma et al. 2015). It is even more so in the digital era; by January 2022, 4.62 billion people, approximately 58.4% of the global population, used online social network platforms, such as Facebook, YouTube, and TikTok (Datareportal 2022). These platforms extend the reach and complexity of our social networks as both friends and strangers online contribute to shaping our opinions and choices. Within such networks, individuals influence and are influenced by others, enabling the diffusion of information and behavior. Platforms that leverage network diffusion can substantially boost their impact and profitability (Shriver et al. 2013). However, understanding network diffusion is complex, involving not just individual preferences but also, the intricate relationships that bind them. This complexity has made network analysis a long-standing research focus (see the book by Jackson 2010).

Network diffusion is studied across disciplines, such as computer science (Kempe et al. 2003, Acemoğlu et al. 2013), economics (Acemoglu et al. 2011, Sadler 2020), operations management (Song and Zipkin 2009, Candogan et al. 2012, Wang and Wang 2017), and epidemiology (Kermack and McKendrick 1927, Drakopoulos and Zheng 2017). In the seminal paper by Kempe et al. (2003), diffusion processes are broadly categorized into progressive and nonprogressive types. Although progressive diffusion deals with unidirectional changes in the state, such as adopting new technology or purchasing a product, our study focuses on *nonprogressive*

diffusion, which allows for bidirectional state transitions. This type is particularly relevant for scenarios in which decisions can be reversed, such as subscription to a membership, belief propagation, or infection dynamics in a pandemic.

The analysis of network diffusion adheres primarily to one of two approaches. The first is microfounded, capturing detailed network topology and stochastic dynamics. Notable models, such as the independent cascade model (Goldenberg et al. 2001) and the linear threshold model (Granovetter 1978, Schelling 1978), provide fine-grained evolution characterization but are computationally intensive. In most cases, simulation happens to be the only viable tool, making optimization, even for a moderate-sized network, time consuming (Chen et al. 2009). In contrast, the second approach offers a macroscopic view, simplifying the diffusion process. For example, Bass models (e.g., Bass 1969) bypass the network structure, focusing on the overall population; others (e.g., Candogan et al. 2012, Jackson et al. 2020) ignore stochasticity and focus on the equilibrium outcome. This macro lens, although sacrificing detailed characterization, enables efficient analysis and sharper insights.

Our work bridges these two approaches in the context of nonprogressive diffusion. We base it on a general microfounded diffusion model that considers heterogeneous agents, local network effects, and network structure. Although analytically characterizing the adoption rate over time in such a model is technically intractable, we propose a fixed-point approximation (FPA) scheme that approximates the dynamics through a set of easily solvable fixed-point equations. Notably, we show that the FPA scheme comes with provable guarantees. Its performance is associated with the network structure and improves for larger and denser networks. We also propose easy-to-calculate metrics at both node and network levels to indicate FPA performance for different network structures. Moreover, the FPA scheme paves the way for optimizing operational decisions in the nonprogressive diffusion context. It enables straightforward problem formulation and algorithm development that are not just computationally efficient but also, yield near-optimal solutions. In summary, through the FPA scheme, we show that the diffusion characterized by a "micromodel" can be accurately approximated by an easy-to-analyze "macromodel," integrating the advantages of both modeling paradigms.

1.1. Contributions and Organization

Our research contributions are summarized as follows.

- Provable approximation scheme for a general nonprogressive diffusion model. We investigate nonprogressive diffusion through a microfounded, dynamic, and stochastic model. Based on this model, we propose the FPA scheme to approximate the adoption probability of each agent. To validate this approach, we develop a nontrivial "fixed-point sandwich" technique, establishing an order-optimal error bound of the FPA scheme in terms of a newly proposed network measure. The bound indicates its superior performance for large and dense networks, which are otherwise challenging to simulate. We further connect the error bound to two novel and easy-to-calculate metrics: the *inverse in-degree centrality* and the *inverse in-degree density*. These metrics, which serve as reliable indicators for FPA performance, provide valuable insight into both node-level and network-wide structures. Our extensive numerical studies further confirm the theoretical results. For real-world networks, it achieves a mean absolute percentage error (MAPE) of less than 3.48% among all instances tested while concurrently accelerating the computation by factors ranging from 70 to 230 compared with the naïve agent-based simulation (ABS).
- Wide applicability in optimizing operational decisions. The FPA scheme offers a powerful tool to reformulate and solve operational decision-making problems in the nonprogressive diffusion setting. Leveraging our approximation error bound, the reformulated problems lead to efficient algorithms that guarantee high-quality decisions, providing a clear advantage over simulation-based methods. We illustrate this with two examples: influence maximization (IM) and optimal (dynamic) pricing. For the IM problem, we show that under technical conditions, the influence function is submodular with regard to the seed set in the reformulated problem. This extends the greedy algorithm to more general settings, significantly improving efficiency. For the optimal pricing problem, the FPA scheme facilitates near-optimal solutions for general static pricing problems, which account for the limiting agent behavior. When perfect price discrimination is implemented, we can reformulate both static and dynamic pricing problems as convex optimization problems under specific conditions, yielding the optimal solution.

The remainder of this paper is structured as follows. In this section, we continue to review the related literature. Section 2 introduces the diffusion model and characterizes the adoption probabilities. In Section 3, we describe the FPA scheme and present our main theoretical results. In Section 4, we establish the order-optimal error bound followed by numerical experiments in Section 5. We apply the FPA scheme to the IM and pricing problems in Section 6. Section 7 concludes this paper.

1.2. Literature Review

Our paper contributes to the network diffusion literature, focusing on both diffusion models and related optimization problems.

1.2.1. Diffusion Models. Various models have been proposed across disciplines to characterize diffusion for specific applications. However, a consistent trade-off can be observed; researchers often have to choose between a comprehensive model and practical efficiency. For instance, the linear threshold (LT) model (Granovetter 1978, Schelling 1978) incorporates the network structure but is computationally challenging as evidenced by Chen et al. (2010). In contrast, the Bass model (Bass 1969) abstracts most of the information about the network structure and individual agents but benefits from analytical tractability, offering closed-form expressions for critical values that facilitate optimization (Agrawal et al. 2021, Lin et al. 2024). Our review categorizes the existing literature into two groups: models that are inherently complex and computationally burdensome, which align with our microfounded model, and models that prioritize analytical tractability, which align with our FPA scheme. Additionally, we discuss how traditional approaches, such as mean-field approximations, have been used to mitigate the computational challenges associated with the first group of models.

Our model builds upon the LT model and is particularly aligned with the nonprogressive LT variant[1] introduced by Kempe et al. (2003). This variant preserves many features of the progressive LT model, but it introduces flexibility by allowing random thresholds to vary independently at each time step in contrast to the fixed random thresholds in the progressive version. We extend the nonprogressive LT model further by introducing agent heterogeneity and accommodating a broader class of randomness distributions, thereby enhancing its applicability to diverse diffusion contexts.

Several parallel streams of work share conceptual connections with our microfounded model. In the field of social learning, agents form beliefs about a binary signal based on their neighbors' beliefs (Jadbabaie et al. 2012, Chandrasekhar et al. 2020, Allon et al. 2021). Although these models emphasize the learning process and the distribution of population beliefs, our work characterizes individual adoptions in arbitrary network structures to facilitate operational decision making. Moreover, our model can be conceptualized as a Markov chain (MC) on graphs, but it fundamentally differs from random walks on the graph (Göbel and Jagers 1974), which track the movement of a single agent across graph vertices. Our model focuses on the diffusion process across a network of agents, requiring an analysis of how the entire agent states evolve collectively over time. Models like Markov random fields also share similarities. For example, the Ising model (Ising 1924) in statistical mechanics describes binary states of vertices influenced by their neighbors, typically aiming to estimate marginal distributions at equilibrium. Unlike these models, our framework formulates a Markov chain to model the dynamic diffusion process, providing a microfoundation for understanding individual agent behaviors by incorporating temporal evolution alongside spatial dependencies.

Another stream of models emphasizes analytical tractability, often at the expense of descriptive power, aligning conceptually with our FPA scheme. For instance, many engineering and economics applications describe interactions using network games (e.g., see Ballester et al. 2006, Candogan et al. 2012, Baron et al. 2022, and Afèche et al. 2023). A central goal of this literature is to analyze various equilibria. Although our FPA scheme is reminiscent of the equilibrium in network games, our focus diverges in its relation to a concrete microfounded model. Similarly, a number of operations management studies incorporate network externality into choice models. This type of work, which serves for subsequent optimization problems, often simplifies the model. For example, some consider only global effects by taking average over the market (Du et al. 2016, Wang and Wang 2017), whereas others use myopic local proxy or specific types of networks (Gopalakrishnan et al. 2023, Xie and Wang 2025). In contrast, our work considers arbitrary network structures in a more general setting.

Finally, our work contributes to the literature on approximation schemes for network diffusion models. Although conceptually related to mean-field approximations (Benaïm and Weibull 2003, Van Mieghem et al. 2009), which offer deterministic descriptions at the population level and focus on equilibrium outcomes, our approach goes further by approximating the dynamic diffusion process at the individual level. In Markov random-field models, mean-field approximations (Yeomans 1992, Boykov et al. 1998) reduce complexity by fitting deterministic approximations into a predefined equilibrium representation. However, our model faces the challenge of evolving correlations among agents' behavior over time, which requires a fundamentally different approach to approximate and analyze. This dynamic consideration not only introduces technical complexities, but also, it enables us to bridge the two distinct streams of models mentioned above. Our FPA scheme facilitates computationally efficient analysis while preserving the key features of the underlying microfounded model, providing a dual perspective for studying and understanding network diffusion processes.

1.2.2. Optimization with Network Diffusion. The FPA scheme has a wide range of applications. In this paper, we highlight its use in two examples—influence maximization and optimal pricing problems. Kempe et al. (2003) first consider the IM problem, which is choosing an influential set of seed agents to maximize total influence, as a discrete optimization problem. They

show that the IM problem, under the LT model, is NP-hard for both progressive and nonprogressive cases. Moreover, evaluating the total influence for different seed sets requires extensive simulations, making it time consuming to obtain even approximate solutions. We refer readers to the survey (Li et al. 2018) for a comprehensive review of existing approaches. These approaches compromise either accuracy or efficiency, and they are not ideal for practical use. With the FPA scheme, we can effectively balance both. For the pricing problem, there is a growing literature in the economics and operations management communities that considers the presence of network effects (Anari et al. 2013, Hu et al. 2020, Li 2020, Huang et al. 2022, Yang and Zhang 2022). Recent studies on the single-item pricing problem with the network effect can be found in Candogan et al. (2012), Du et al. (2018), and Nosrat et al. (2021). Compared with others, our proposed framework can be used in a more general setting. Moreover, much of this literature focuses on the static pricing problem, aiming to determine an equilibrium price when considering network effects. In this work, we also address the dynamic pricing problem under certain conditions, which allows us to incorporate transient behaviors before reaching equilibrium.

2. Nonprogressive Network Diffusion Model

In this section, we first introduce the nonprogressive diffusion model and then, characterize the limiting behavior of each agent. Although this model can be applied to various nonprogressive diffusion settings, we use service adoption on an online social network platform for illustration.

2.1. Preliminaries and Formulation

We model the social network platform (e.g., TikTok) as a graph $G = (V, E)$ with n agents, where $V := \{1, 2, \ldots, n\}$ is the set of agents and $E := \{1, 2..., |E|\}$ is the set of directed edges. A directed edge $(i, j) \in E$, where $i, j \in V$, implies that agent j is influenced by agent i, and we call i an *in-neighbor* of j. We interpret $(i, j) \in E$ as j following i on the platform. We use $\mathcal{N}_i$ to denote the set of all in-neighbors for agent i (i.e., $\mathcal{N}_i := \{j \in V : (j, i) \in E\}$) and $d_i := |\mathcal{N}_i|$ to denote the in-degree (i.e., the number of in-neighbors). Throughout, we use agent and node interchangeably.

We use t to denote the discrete time period, starting with $t = 0$ as the service launch time. Define $Y_i(t) \in \{0, 1\}$ as the state of agent i at time t, where $Y_i(t) = 1$ (respectively, $Y_i(t) = 0$) means adoption (respectively, nonadoption) of the service in this period. The initial state $\mathbf{Y}(0)$ follows an arbitrary distribution on $\{0, 1\}^n$. For all $t \geq 1$, each agent i decides whether to adopt the service based on their realized utility $u_i(t)$ during that period, which is given by

$$u_i(t) := v_i + \beta \cdot \frac{\sum_{j \in \mathcal{N}_i} Y_j(t-1)}{d_i} + \epsilon_i(t). \tag{1}$$

Without loss of generality, we normalize the utility of nonadoption to zero so that $Y_i(t) = \mathbb{1}\{u_i(t) \geq 0\}$. As shown in (1), $u_i(t)$ consists of three parts: the idiosyncratic intrinsic value v_i, *the local network effect* $\beta \cdot \sum_{j \in \mathcal{N}_i} Y_j(t-1)/d_i$, and random noise $\epsilon_i(t)$. The value v_i reveals the personalized preference and remains constant over time. Analytically, it can be estimated from features, such as demographics and behavioral data, with the support of big data. It may also be affected by the platform strategies. For example, the price of a paid service (e.g., YouTube Premium) will definitely affect whether and how the agent likes it. The local network effect term captures peer influence, with β quantifying the network effect intensity. If agent i has no in-neighbors (i.e., $\mathcal{N}_i = \emptyset$), we set this term to zero. Finally, we assume that random noise $\epsilon_i(t)$ is independent and identically distributed (i.i.d.) across agents and time, with $\mathbb{E}[\epsilon_i(t)] = 0$. Although the i.i.d. assumption is standard, we acknowledge that real-world scenarios may introduce more complexity. For example, agents may gradually reveal their preferences, leading to a diminishing variance in random noise over time. Alternatively, public news may introduce positively correlated noise across agents at the same time. Despite these complexities, the approximation approach proposed later in this work remains valid, although the approximation error may increase. Online Appendix A.2 provides a detailed discussion on relaxing the i.i.d. assumption. For now, we impose no further constraints on its distribution, except for the following mild condition.

Assumption 1 (Lipschitz Continuity). *The random noise $\epsilon_i(t)$ has an L-Lipschitz continuous cumulative distribution function (CDF): $|F_\epsilon(x) - F_\epsilon(y)| \leq L|x - y|$ for any $x, y \in \mathbb{R}$.*

This assumption ensures a sufficiently smooth noise distribution. Assumption 1 holds for any continuous distribution with a bounded probability density function (PDF), such as uniform, logistic, or normal distribution, making them compatible with our model. The parameter L characterizes the sensitivity of the agent's behavior to random noise. We note that our result remains valid even if L is not defined across the entire CDF as in Assumption 1. For instance, for a given intrinsic value $\mathbf{v}$, one can define an instance-dependent L that need not be the Lipschitz constant over the entire real line but only within the domain $[-\max_{i \in V} v_i - \beta, -\min_{i \in V} v_i]$. In this case, our results still hold. However, for clarity and consistency, we maintain Assumption 1 throughout.

A natural goal of this model is to quantify the total diffusion in the network. In line with previous studies

(Kempe et al. 2003), we focus on the limiting adoption probability.[2] Provided that it converges, this also represents the accumulated reward (frequency of adoptions) in the long run. However, to discuss such behavior in a meaningful way, we must first ensure its existence. Without additional assumptions, the process can diverge or yield multiple outcomes, complicating the focus on a single outcome of interest. To guarantee a unique outcome, we impose Assumption 2, which excludes divergent or periodic behavior (Proposition 1) and ensures the existence of a valid fixed-point solution for our approximation (Proposition 2).

Assumption 2 (Bounded Network Effects). *The network effect intensity satisfies* $|\beta| < 1/L$.

This assumption ensures that agents' behavior is not overly sensitive to network effects or random noise. It is well established in the literature and is typically satisfied in practice. Specifically, it aligns with conditions in the network diffusion literature (Kempe et al. 2003) and assumptions in social economics research (e.g., see Horst and Scheinkman 2006, Wang and Wang 2017, and Jackson et al. 2020). Studies that empirically calibrate β in real-world applications further support the validity of this assumption. For instance, Wang and Wang (2017) assume that random noise follows a standard logistic distribution, logistic$(0,1)$[3] (logit model), resulting in $L = 1/4$, with calibrated β ranging from 1.391 to 2.392. Similarly, Chen et al. (2021) assume a standard normal distribution (probit model), yielding $L = 1/\sqrt{2\pi}$, with calibrated β ranging from 0.69 to 1.11. These empirical results provide strong evidence that the assumption holds in practice. Although our theoretical analysis is based on Assumption 2, we acknowledge that it may not always hold. Therefore, Online Appendix D.1.1 extends our discussion with numerical experiments that explore the implications of violating Assumption 2. For the rest of our theoretical analysis, we assume positive network effects (i.e., $0 < \beta < 1/L$). However, our results can also be generalized to account for negative network effects, where $-1/L < \beta < 0$.

We remark on the notation. Hereafter, we use bold math notation to denote the collection of a particular variable over all agents in vector form. We represent a specific diffusion instance by a quadruple $(G, \mathbf{v}, F_\epsilon(\cdot), \beta)$ because the network structure and intrinsic values identify a diffusion case. Meanwhile, the noise distribution and the network effect intensity constitute the diffusion environment. Sequences of such instances are represented by a series of these quadruples.

2.2. A Markov Chain Perspective

Notably, each diffusion instance can be characterized by a Markov chain, of which the state space is the set of indicator vectors denoting all possible combinations of adoption decisions, represented by $\{0,1\}^n$. The transition probability from state $\mathbf{y}$ to $\mathbf{y}'$ can be computed as

$$\begin{aligned} P(\mathbf{y}, \mathbf{y}') &= \prod_{i \in V} \mathbb{P}(Y_i(t) = y_i' \mid \mathbf{Y}(t-1) = \mathbf{y}) \\ &= \prod_{i \in V} F_\epsilon\left(-v_i - \beta \frac{\sum_{j \in \mathcal{N}_i} y_j}{d_i}\right)^{1-y_i'} \\ &\quad \cdot \left[1 - F_\epsilon\left(-v_i - \beta \frac{\sum_{j \in \mathcal{N}_i} y_j}{d_i}\right)\right]^{y_i'}. \end{aligned}$$

Our primary interest lies not in the individual MC states but rather, in the overall adoption probability for each agent. To that end, we define the adoption probability of agent i at time t as

$$q_i(t) := \mathbb{P}(Y_i(t) = 1) = \sum_{\mathbf{y} \in \{0,1\}^n} \mathbb{1}\{y_i = 1\} \cdot \mathbb{P}(\mathbf{Y}(t) = \mathbf{y}). \quad (2)$$

We have the following proposition on the limiting behavior of $\mathbf{q}(t)$ when t tends to infinity.

Proposition 1 (Limiting Adoption Probability). *Under Assumptions 1 and 2, for any initial state* $\mathbf{Y}(0) \in \{0,1\}^n$, *the adoption probability of each agent i converges to*

$$\lim_{t \to \infty} q_i(t) = q_i^* := \sum_{\mathbf{y} \in \{0,1\}^n} \mathbb{1}\{y_i = 1\} \cdot \pi(\mathbf{y})$$

as t tends to infinity, where $\boldsymbol{\pi}$ *is the stationary distribution of the MC that satisfies* $\boldsymbol{\pi} = \boldsymbol{\pi} P$.

As shown in the proof of Proposition 1 (Online Appendix A.1), this MC has a *single aperiodic recurrent class*, ensuring the existence of the limiting distribution $\boldsymbol{\pi}$ and hence, $\mathbf{q}^*$. By standard MC theory,

$$\begin{aligned} &\lim_{t \to \infty} \frac{1}{t} \cdot \sum_{\tau=1}^{t} Y_i(\tau) = q_i^* \quad \text{a.s.} \quad \text{and} \\ &\lim_{t \to \infty} \frac{1}{t} \cdot \sum_{\tau=1}^{t} q_i(\tau) = q_i^*, \quad \forall i \in V, \end{aligned} \quad (3)$$

for any initial state $\mathbf{Y}(0)$. As a result, various operational problems, such as the IM (Section 6.1) and optimal pricing problem (Section 6.2), can be generally framed as

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{maximize}} \quad g(\mathbf{q}^*(G, \mathbf{v}(\mathbf{x}), F_\epsilon(\cdot), \beta), \mathbf{x}). \quad (4)$$

Here, $\mathbf{x}$ represents the platform decisions, with $\mathcal{X}$ denoting its feasible set. For simplicity, we only consider decisions that influence diffusion outcomes by altering intrinsic values. With a slight abuse of notation, $\mathbf{v}(\cdot)$ represents intrinsic values as a function of platform decisions, and $\mathbf{q}^*(\cdot)$ denotes the mapping from a diffusion instance to the limiting probability vector. Finally, $g(\cdot,\cdot)$ is the objective function that depends on the diffusion outcome $\mathbf{q}^*$ and decision $\mathbf{x}$. For example, the IM problem can be formulated as (4), where the decision $\mathbf{x}$

is to set the intrinsic utility of a set of seed users at sufficiently high levels and the objective $g(\cdot,\cdot)$ is the expected total limiting adoptions, $\sum_{i\in V} q_i^*$. For the optimal pricing problem, the decision $\mathbf{x}$ is the price vector that affects the intrinsic value of each agent, and the objective $g(\cdot,\cdot)$ is the expected profit under the limiting adoption probability (i.e., $\sum_{i\in V} q_i^* x_i$). The specific formulations of these problems will be presented in Section 6.

Solving Problem (4) is challenging because of the absence of closed-form expressions for $\mathbf{q}^*$ and the exponential growth of the MC states. It is intractable to construct the transition matrix even for a moderately sized network, let alone to calculate $\mathbf{q}^*$. Therefore, Problem (4) is generally intractable analytically, which motivates us to develop our approximation scheme presented in Section 3.

3. Fixed-Point Approximation Scheme

In this section, we introduce the FPA scheme and provide a comprehensive analysis of its performance, including upper bounds for approximation errors. Additionally, we also present two easy-to-compute metrics designed to evaluate FPA performance; see (6).

3.1. Overview and Motivating Example

For a given diffusion instance $(G,\mathbf{v},F_\epsilon(\cdot),\beta)$, we will show that the limiting adoption probability $\mathbf{q}^*$ can be reasonably approximated by the solution $\boldsymbol{\mu}^*$ of the following system of equations:

$$\mu_i = 1 - F_\epsilon\left(-v_i - \beta\frac{\sum_{j\in\mathcal{N}_i}\mu_j}{d_i}\right) \text{ for all } i\in V. \qquad (5)$$

Although our main result focuses on the limiting adoption probability, the full analysis goes beyond this. Specifically, it captures the entire diffusion process, including transient behaviors. Because (5) represents the limiting state of a fixed-point iteration, which is a deterministic dynamical system, the FPA scheme can approximate the entire diffusion process with the evolution of this system. As such, the FPA scheme is also applicable to network diffusion problems with transient behaviors.

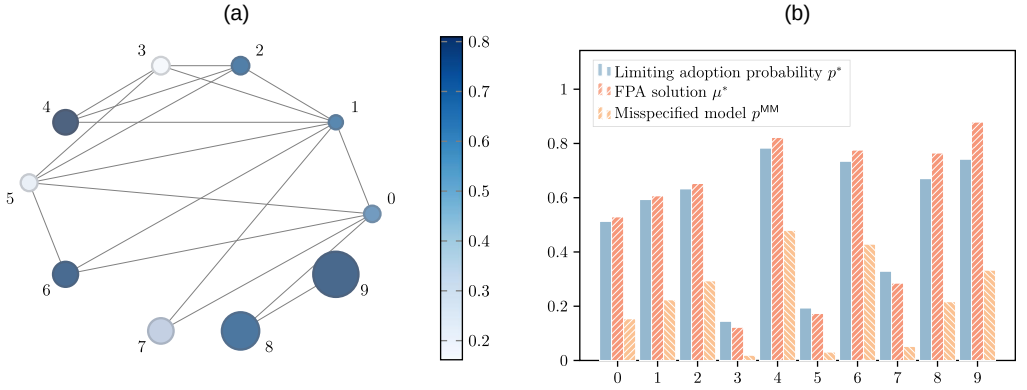
We begin with a motivating example to illustrate the quality of the FPA scheme as shown in Figure 1. For detailed information on this example instance, including numerical results, please refer to Online Appendix D.1. In Figure 1, to provide an intuitive understanding of the network effects, we also introduce a misspecified model (MM) as a benchmark. For the MM, the adoption probability for agent i is calculated as $q_i^{\mathrm{MM}} := \mathbb{E}[\mathbb{1}\{v_i+\epsilon_i\geq 0\}]$, ignoring network effects. Figure 1(a) presents the network structure and approximation results for the example. Clearly, nodes with fewer neighbors exhibit larger errors, whereas well-connected nodes yield smaller errors. Figure 1(b) compares the values of $\mathbf{q}^*$, $\boldsymbol{\mu}^*$, and $\mathbf{q}^{\mathrm{MM}}$, showing the impact of network effects as evidenced by the discrepancy between $\mathbf{q}^*$ and $\mathbf{q}^{\mathrm{MM}}$. Using $\mathbf{q}^*$ as the baseline, the mean absolute errors for $\boldsymbol{\mu}^*$ and $\mathbf{q}^{\mathrm{MM}}$ are 0.045 and 0.310, respectively. These findings confirm the feasibility of the FPA scheme and suggest that the approximation is more accurate for agents with central positions in the network.

Next, we theoretically analyze the deviation between $\mathbf{q}^*$ and $\boldsymbol{\mu}^*$, focusing on how it depends on network structure. The key challenge lies in the temporal and spatial correlations in adoptions, which are further complicated by the nonlinear evolution introduced by the general noise distribution.

3.2. The Approximation Error

We first remark on the notation before formally analyzing the error bound. Given a network $G=(V,E)$, we define the matrix $\tilde{\mathbf{A}}\in\mathbb{R}^{n\times n}$ such that $\tilde{A}_{ij}=1/d_i$ if an edge is directed from j to i and $\tilde{A}_{ij}=0$ otherwise. This matrix is a scaled version of the adjacency matrix $\mathbf{A}$, where $A_{ij}=1$ if there is an edge directed from i to j

Figure 1. (Color online) A 10-Node Example to Illustrate the FPA Scheme



Notes. In panel (a), for each agent i, the color denotes the true values of q_i^*, and the size denotes the absolute error $|q_i^*-\mu_i^*|$. (a) Approximation results on the 10-node example instance. (b) Values of adoption probabilities.

and $A_{ij} = 0$ otherwise. One obtains $\tilde{\mathbf{A}}$ by scaling row i of $\mathbf{A}^\top$ by $1/d_i$. Notably, $\tilde{\mathbf{A}}$ is a row-stochastic matrix, meaning that $\tilde{\mathbf{A}}\mathbf{e} = \mathbf{e}$, where $\mathbf{e}$ is a vector of ones. We also introduce the vector $\mathbf{b} := (1/d_1, 1/d_2, \ldots, 1/d_n)^\top$, which contains the reciprocals of each node's in-degree. Lastly, we define $\rho := L\beta$ and let $d_{\min}$ be the minimum in-degree of the network, with $d_{\min} > 0$.

We further introduce two metrics in our analysis, which we term the *inverse in-degree centrality* $\mathcal{C}(G;\rho)$ and the *inverse in-degree density* $\mathcal{D}(G)$. They are defined as follows:

$$\mathcal{C}(G;\rho) := (1-\rho)\left(\mathbf{I} + \sum_{\ell=1}^{\infty} \rho^\ell \tilde{\mathbf{A}}^\ell\right)\mathbf{b} = (1-\rho)(\mathbf{I} - \rho\tilde{\mathbf{A}})^{-1}\mathbf{b}$$

$$\text{and} \quad \mathcal{D}(G) := \frac{\mathbf{e}^\top \mathbf{b}}{n}. \tag{6}$$

The *inverse in-degree centrality* $\mathcal{C}(G;\rho)$ is an n-dimensional vector that captures the centrality of each agent, with its ith entry denoted by $\mathcal{C}_i(G;\rho)$. It bears similarities with the classical Bonacich centrality (Bonacich 1987), taking the form of a Neumann series. However, by incorporating the inverse of in-degrees, it is tailored to evaluate the FPA performance. This centrality is well defined because $\tilde{\mathbf{A}}$ is row stochastic and $\rho < 1$ by Assumption 2. The *inverse in-degree density* $\mathcal{D}(G)$ is a scalar that represents the average inverse in-degree of all agents, serving as an aggregate measure of FPA performance. Together, these two metrics offer both individualized and holistic views on network structure. This dual perspective not only enhances our understanding of the FPA scheme but also, provides actionable insights into its application across different network configurations.

For clarity, we adopt the subscript $_{\text{ew}}$ to represent entry-wise operations on vectors. For instance, for vector $\mathbf{q}$, we define $|\mathbf{q}|_{\text{ew}} := (|q_1|, |q_2|, \ldots, |q_n|)^\top$ and $\mathbf{q}^{\frac{1}{2}}_{\text{ew}} := (\sqrt{q_1}, \sqrt{q_2}, \ldots, \sqrt{q_n})^\top$. Let $\odot$ denote the Hadamard product, which is essentially a component-wise product of two matrices. Additionally, we also define the constant $C_\rho := \rho/(1-\rho)^{3/2}$. We now present our key technical result in Theorem 1.

Theorem 1 (Entry-Wise Error Bound of the FPA Scheme). *Under Assumptions 1 and 2, for any diffusion instance* $(G, \mathbf{v}, F_\epsilon(\cdot), \beta)$, *the absolute difference between the limiting adoption probability* $\mathbf{q}^*$ *and the fixed-point solution* $\boldsymbol{\mu}^*$ *can be upper bounded by*

$$|\mathbf{q}^* - \boldsymbol{\mu}^*|_{\text{ew}} \le \rho(\mathbf{I} - \rho\tilde{\mathbf{A}})^{-1}[\mathcal{B}(G;\rho)]^{\frac{1}{2}}_{\text{ew}} \le C_\rho \cdot [\mathcal{C}(G;\rho)]^{\frac{1}{2}}_{\text{ew}},$$

where $\mathcal{B}(G;\rho) := \left(\sum_{\tau=1}^{\infty} \rho^{2\tau-2}\tilde{\mathbf{A}}^\tau \odot \tilde{\mathbf{A}}^\tau\right)\mathbf{e}$.

Theorem 1 characterizes the entry-wise error between the limiting probability $\mathbf{q}^*$ and our FPA solution $\boldsymbol{\mu}^*$. We show that this error is closely related to the network structure. Specifically, it is linked to the inverse in-degree centrality $\mathcal{C}(G;\rho)$ via an intermediate measure $\mathcal{B}(G;\rho)$. Although $\mathcal{B}(G;\rho)$ involves an infinite sum and cannot be computed efficiently, it offers a detailed characterization of how the error propagates across the network. Moreover, it plays a key role in establishing the order-optimal results discussed later in Section 4.2. In contrast, the inverse in-degree centrality $\mathcal{C}(G;\rho)$ offers a more concise and interpretable metric to quantify the error. Each component of $\mathcal{C}(G;\rho)$ can be viewed as a weighted sum of entries in $\mathbf{b}$, the vector of inverse in-degrees. These weights reflect the connectivity between nodes. Specifically, the weight associated with inverse in-degree b_j in the ith term of $\mathcal{C}(G;\rho)$ is $(1-\rho)\sum_{P\in\mathcal{P}_{(j,i)}}\rho^{|P|}\prod_{k\in P}1/d_k$, where $\mathcal{P}_{(j,i)}$ denotes the set of directed paths from agent j to i. Remarkably, this weight decays exponentially with the length of the directed path. As a result, the inverse in-degree centrality of each node is predominantly affected by the inverse in-degrees of its nearby nodes. This result rationalizes the observations in Figure 1.

To further understand the intuition of $\mathcal{C}(G;\rho)$, we let $d_{\min,i}(\tau)$ be the minimum in-degree of any node j that is connected to node i via a path of length τ. It is straightforward that $d_{\min,i}(\tau) \ge d_{\min}$. By expanding the definition of inverse in-degree centrality, it then holds that

$$\mathcal{C}_i(G;\rho) \le (1-\rho)\sum_{\tau=0}^{\infty}\frac{\rho^\tau}{d_{\min,i}(\tau)} \le \frac{1}{d_{\min}} \text{ for all } i \in V, \tag{7}$$

where the first inequality holds because $\tilde{\mathbf{A}}$ is a row-stochastic matrix. Consequently, we arrive at

$$|q_i^* - \mu_i^*| \le C_\rho \cdot \sqrt{(1-\rho)\sum_{\tau=0}^{\infty}\frac{\rho^\tau}{d_{\min,i}(\tau)}},$$

which suggests that the error is small for nodes that have both large in-degrees and are more distant from nodes with low in-degrees. Hence, we can establish the subsequent corollary.

Corollary 1 (ℓ_∞-Norm Error Bound). *Under Assumptions 1 and 2, for any diffusion instance* $(G, \mathbf{v}, F_\epsilon(\cdot), \beta)$, *the* ℓ_∞*-norm of the difference between* $\mathbf{q}^*$ *and* $\boldsymbol{\mu}^*$ *can be upper bounded by*

$$\|\mathbf{q}^* - \boldsymbol{\mu}^*\|_\infty \le C_\rho \cdot \sqrt{\frac{1}{d_{\min}}}. \tag{8}$$

Corollary 1 removes the dependence on the specific network structure from the bound to highlight a worst-case convergence rate as the network expands. Specifically, for a sequence of diffusion instances characterized by an increasing minimum in-degree $d_{\min}$, the maximal deviation shrinks at a rate of $\mathcal{O}(\sqrt{1/d_{\min}})$. As $d_{\min}$ approaches infinity, $\boldsymbol{\mu}^*$ is asymptotically equal to $\mathbf{q}^*$. This simplified bound clearly indicates that the FPA scheme can perform better in larger and denser networks. We remark that although the ℓ_∞-norm error bound is intuitively appealing, it can be much looser than the entry-wise bound presented in Theorem 1. It

relies solely on the minimal in-degree $d_{\min}$, making it overly conservative and sensitive to isolated outliers. In most real-world networks, the minimal in-degree $d_{\min}$ is often quite small even if its size n is large, limiting the applicability of this bound.

Corollary 2 below addresses this limitation by introducing a bound based on the scaled ℓ_1-norm. Define $r(G) := \max_{i\in V}\left(\sum_{j=1}^{n} A_{ij}/\sum_{j=1}^{n} A_{ji}\right)$ as the largest ratio between out-degree and in-degree.

Corollary 2 (Scaled ℓ_1-Norm Error Bound). *Under Assumptions 1 and 2, for any diffusion instance $(G,\mathbf{v},F_\epsilon(\cdot),\beta)$, the scaled ℓ_1-norm of the difference can be upper bounded by*

$$\frac{1}{n}\|\mathbf{q}^* - \boldsymbol{\mu}^*\|_1 \le C_\rho \cdot \sqrt{\frac{\|\mathcal{C}(G;\rho)\|_1}{n}}. \tag{9}$$

If $r(G) < 1/\rho$, the bound can be further simplified as

$$\frac{1}{n}\|\mathbf{q}^* - \boldsymbol{\mu}^*\|_1 \le C_\rho \cdot \sqrt{\frac{1-\rho}{1-\rho r(G)}\cdot \mathcal{D}(G)}. \tag{10}$$

In light of (7), the network-structure-free bound on $\frac{1}{n}\|\mathbf{q}^* - \boldsymbol{\mu}^*\|_1$ is also of the order $\mathcal{O}(\sqrt{1/d_{\min}})$. However, Corollary 2 provides more meaningful bounds. Specifically, (9) bounds the scaled ℓ_1-norm of the error by that of the inverse in-degree centrality $\mathcal{C}(G;\rho)$. More transparent results can be obtained when considering the largest out-degree-to-in-degree ratio $r(G)$. A smaller value of $r(G)$ indicates a more evenly distributed degree structure, with $r(G) \ge 1$ always holding. In poorly conditioned networks, where $r(G)$ is large, substantial weights may apply to nodes with a small in-degree, making the bound approach the worst-case rate $\mathcal{O}(\sqrt{1/d_{\min}})$. In contrast, for well-conditioned networks where $r(G) < 1/\rho$, we obtain a bound in (10), which is characterized by the inverse in-degree density $\mathcal{D}(G)$. Unlike $d_{\min}$, which focuses on extreme nodes, $\mathcal{D}(G)$ offers a holistic view by capturing the average characteristics of the network. Additionally, $\mathcal{D}(G)$ is also computationally more efficient than both $\mathcal{B}(G;\rho)$ and the inverse in-degree centrality $\mathcal{C}(G;\rho)$. As a consequence, $\mathcal{D}(G)$ serves as a more practical and efficient indicator for FPA performance across different networks.

Looking more closely at the impact of $r(G)$, we find that the upper bound becomes tighter as $r(G)$ decreases, indicating better FPA performance in more balanced networks. We highlight that the assumption $r(G) < 1/\rho$ for the second part of Corollary 2 is not restrictive in general. Notably, all undirected graphs and balanced directed graphs satisfy this. Studies, such as Mislove (2009), also validate the balanced nature of social networks in practice. In particular, active agents (i.e., those who create many links) also tend to be popular (i.e., they are the target of many links). This high correlation is generally attributed to the prevalence of reciprocal links in social networks.

We remark on three facts. First, all of the aforementioned bounds apply to networks where $d_{\min} > 0$. For any stand-alone node i with no in-neighbors, the network effect term in (1) is zero, meaning that $\mu_i^* = q_i^*$ trivially holds. Including such nodes only tightens the derived bounds. Second, the constant C_ρ decreases with ρ and converges to zero as ρ approaches zero. Therefore, our bounds suggest that FPA works better when ρ, which represents the compound effect of network externality and random noise, is small. As our theoretical analysis is based on Assumption 2, when the assumption is violated, the bound may lose validity and diverge to infinity. However, in such cases, a constant, such as one or a problem-specific value, can serve as a trivial bound. We defer the investigation of scenarios when this assumption does not hold in Online Appendix D.1.1. Third, we can also integrate the FPA scheme with simulation techniques, balancing between precision and computational efficiency. In Section 5.4, we propose such a mixture scheme and empirically demonstrate its effectiveness. By selectively resampling low-degree agents, the mixture scheme enhances accuracy with minimal computational overhead, offering scalable, high-precision solutions for different networks.

On the operational side, the importance of the FPA solution $\boldsymbol{\mu}^*$ lies in the fact that it allows us to reformulate and simplify Problem (4). Instead of solving (4) directly, we can replace $\mathbf{q}^*$ with $\boldsymbol{\mu}^*$ and approximate Problem (4) as follows:

$$\underset{\mathbf{x}\in\mathcal{X},\,\boldsymbol{\mu}}{\text{maximize}} \quad g(\boldsymbol{\mu},\mathbf{x}) \quad \text{s.t. } \boldsymbol{\mu} = \mathbf{h}(\boldsymbol{\mu}\,;\, G,\mathbf{v}(\mathbf{x}),F_\epsilon(\cdot),\beta), \tag{11}$$

where $\mathbf{h}(\cdot\,;G,\mathbf{v},F_\epsilon(\cdot),\beta)$ is the adoption evolution operator (AEO) induced by the diffusion instance $(G,\mathbf{v},F_\epsilon(\cdot),\beta)$, which we will formally define using (12) in Section 3.3. The approximate Problem (11) offers an explicit formulation by incorporating the FPA scheme as a constraint. This stands in contrast to the implicit variable $\mathbf{q}^*$ in (4), which emerges from a complex stochastic process.

We advocate for the approximate Problem (11) for three reasons: (i) *theoretical guarantee*, (ii) *computational efficiency*, and (iii) *closed-form expression*. The first reason is *theoretical guarantee*. The FPA scheme is particularly effective for large, dense networks, providing strong theoretical guarantees. In reality, many real-world networks are large and expanding, making the FPA scheme a promising tool (see Section 5.3). The second reason is *computational efficiency*. The FPA scheme outperforms agent-based simulation in computational efficiency as the fixed-point iteration converges to the FPA solution in linear time (Rheinboldt 1998). The third reason is *closed-form expression*. Problem (11) is more tractable than (4), enabling more efficient algorithms tailored for specific problems (see Section 6).

3.3. Proof Sketch of Theorem 1

In this section, we sketch the proof of Theorem 1, which is our main methodological contribution. The key idea is to construct a deterministic process $\{\boldsymbol{\mu}(t)\}_{t=0}^{\infty}$ for a given instance $(G,\mathbf{v},F_\epsilon(\cdot),\beta)$. We show that $\{\boldsymbol{\mu}(t)\}_{t=0}^{\infty}$ closely aligns with the evolution of the adoption probability $\{\mathbf{q}(t)\}_{t=0}^{\infty}$.

Specifically, we define $\{\boldsymbol{\mu}(t)\}_{t=0}^{\infty}$ as a deterministic dynamic system throughout the time horizon:

$$\mu_i(t)=\begin{cases} q_i(0) & t=0 \\ 1-F_\epsilon\left(-v_i-\beta\dfrac{\sum_{j\in\mathcal{N}_i}\mu_j(t-1)}{d_i}\right) & t>0 \end{cases} \quad \text{for all } i\in V. \tag{12}$$

Without loss of generality, we assume $\mathbf{Y}(0)=\mathbf{0}$, so $\mathbf{q}(0)=\mathbf{0}$. By Proposition 1, $\mathbf{q}^*$ is unique regardless of the initial distribution $\mathbf{Y}(0)$, and therefore, any error bound derived under $\mathbf{Y}(0)=\mathbf{0}$ applies to arbitrary initial distributions of $\mathbf{Y}(0)$. We define $\mathbf{h}:\mathbb{R}^n\to\mathbb{R}^n$ as the mapping function that expresses $\{\boldsymbol{\mu}(t)\}_{t=0}^{\infty}$ in the form $\boldsymbol{\mu}(t)=\mathbf{h}(\boldsymbol{\mu}(t-1))$ for $t\geq 1$. We refer to $\mathbf{h}(\cdot)$ as an adoption evolution operator and introduce a family of auxiliary AEOs: $\mathcal{H}:=\{\mathbf{h}_{\boldsymbol{\zeta}}(\cdot)=\mathbf{h}(\cdot)+\boldsymbol{\zeta}:\boldsymbol{\zeta}\in\mathbb{R}^n\}$. We proceed by discussing the properties of any AEO $\mathbf{h}\in\mathcal{H}$ and its role in shaping the dynamic system $\{\boldsymbol{\mu}(t)\}_{t=0}^{\infty}$.

Proposition 2 (Partial Order Preserving, Existence, and Uniqueness). *Any AEO $\mathbf{h}\in\mathcal{H}$ satisfies the following properties* (i) *and* (ii), *and the induced dynamic system $\{\boldsymbol{\nu}(t)\}_{t=0}^{\infty}$ defined by fixed-point iteration $\boldsymbol{\nu}(t)=\mathbf{h}(\boldsymbol{\nu}(t-1))$ satisfies the following property* (iii):

i. $\mathbf{h}(\mathbf{a})\leq\mathbf{h}(\mathbf{b})$ *if* $\mathbf{a}\leq\mathbf{b}$;

ii. *there exists a unique fixed-point solution $\boldsymbol{\nu}^*\in\mathbb{R}^n$ with $\mathbf{h}(\boldsymbol{\nu}^*)=\boldsymbol{\nu}^*$; and*

iii. *for any initial state $\boldsymbol{\nu}(0)$, the dynamic system $\{\boldsymbol{\nu}(t)\}_{t=0}^{\infty}$ satisfies $\lim_{t\to\infty}\boldsymbol{\nu}(t)=\boldsymbol{\nu}^*$.*

Property (i) of Proposition 2 follows directly from the definition of $\mathbf{h}(\cdot)$, whereas properties (ii) and (iii) of Proposition 2 follow because $\mathbf{h}(\cdot)$ is a contraction mapping. Note that $\{\boldsymbol{\mu}(t)\}_{t=0}^{\infty}$ is a special case of the induced dynamic system $\{\boldsymbol{\nu}(t)\}_{t=0}^{\infty}$. Therefore, for any diffusion instance under Assumptions 1 and 2, we can always find a well-defined FPA solution $\boldsymbol{\mu}^*$ for limiting adoption probability $\mathbf{q}^*$ by solving the system of equations $\mathbf{h}(\boldsymbol{\mu})=\boldsymbol{\mu}$.

In this proof, we, in fact, show a claim stronger than Theorem 1; $\{\boldsymbol{\mu}(t)\}_{t=0}^{\infty}$ and $\{\mathbf{q}(t)\}_{t=0}^{\infty}$ are always uniformly close to each other. Taking the limit as $t\to\infty$, we then show Theorem 1. For this, we face two challenges. The first is the temporal and spatial dependencies in adoptions. An agent's adoption utility is directly shaped by their in-neighbors, and these localized correlations accumulate and propagate across the network over time. The second challenge arises from the nonlinearity of the CDF F_ϵ of a general distribution, which complicates the analytical tracking of the evolution of adoption states, particularly in characterizing adoption correlations. To address these, our subsequent efforts focus on bounding the spatiotemporal variances and the nonlinear dynamics sequentially.

First, we focus on the local network effect term in (1), which represents an average over a set of mutually dependent random variables. We characterize the covariance matrix of adoption states $\mathbf{Y}(t)$ in Lemma 1, and then, we relate it to the variance of the local network effect term in Lemma 2.

Lemma 1 (Covariance Matrix). *Under Assumptions 1 and 2, for any diffusion instance $(G,\mathbf{v},F_\epsilon(\cdot),\beta)$ and $t\geq 1$, the covariance matrix of adoption states $\mathbf{Y}(t)$ can be upper bounded by*

$$\boldsymbol{\Sigma}(t)\leq\frac{1}{4}\left[\mathbf{I}+\sum_{\tau=1}^{t-1}\rho^{2\tau}\tilde{\mathbf{A}}^{\tau}(\tilde{\mathbf{A}}^{\top})^{\tau}\right].$$

The proof of Lemma 1 establishes the upper bound on $\boldsymbol{\Sigma}(t)$ through induction. A key step is demonstrating that adoption states $\mathbf{Y}(t)$ remain positively associated (see Definition EC.2 in Online Appendix B.2), which is a stronger form of positive correlation. This property enables the derivation of an iterative upper bound on the covariance matrix, despite the nonlinear dynamics.

To quantify the variance of the local network effect, we introduce $\kappa_i(t):=\mathrm{Var}(\sum_{j\in\mathcal{N}_i}Y_j(t)/d_i)$. Lemma 2 then provides an upper bound on this variance for each period based on Lemma 1.

Lemma 2 (Variance of Network Effect). *Under Assumptions 1 and 2, for any diffusion instance $(G,\mathbf{v},F_\epsilon(\cdot),\beta)$ and $t\geq 1$, the variance of the local network effect can be upper bounded by*

$$\boldsymbol{\kappa}(t)\leq\frac{1}{4}\left(\sum_{\tau=1}^{t}\rho^{2\tau-2}\tilde{\mathbf{A}}^{\tau}\odot\tilde{\mathbf{A}}^{\tau}\right)\mathbf{e}\leq\frac{1}{4}\left(\mathbf{I}+\sum_{\tau=1}^{t-1}\rho^{2\tau}\tilde{\mathbf{A}}^{\tau}\right)\mathbf{b}.$$

In Lemma 2, the first upper bound provides a detailed measure that characterizes the evolution of the variance $\boldsymbol{\kappa}(t)$. However, this measure is difficult to compute in general. To improve both interpretability and computational efficiency, we introduce the second looser bound, which is particularly useful in the asymptotic case. Specifically, for $t\geq 1$, we can bound $\boldsymbol{\kappa}(t)$ as follows:

$$\begin{aligned}\boldsymbol{\kappa}(t)&\leq\frac{1}{4}\left(\sum_{\tau=1}^{\infty}\rho^{2\tau-2}\tilde{\mathbf{A}}^{\tau}\odot\tilde{\mathbf{A}}^{\tau}\right)\mathbf{e}=\frac{1}{4}\mathcal{B}(G;\rho)\\&\leq\frac{1}{4}\left[\mathbf{I}+\sum_{\tau=1}^{\infty}\rho^{2\tau}\tilde{\mathbf{A}}^{\tau}\right]\mathbf{b}=\frac{1}{4}(\mathbf{I}-\rho^2\tilde{\mathbf{A}})^{-1}\mathbf{b}.\end{aligned} \tag{13}$$

Here, we highlight both $(1/4)\cdot\mathcal{B}(G;\rho)$ and $(1/4)\cdot(\mathbf{I}-\rho^2\tilde{\mathbf{A}})^{-1}\mathbf{b}$ as upper bounds. In the special case where the underlying network structure is a directed ring, both expressions collapse into the same bound.

However, in general, each bound offers unique insight and has its own advantages.

On one hand, $(1/4)\cdot\mathcal{B}(G;\rho)$ precisely captures how variation in the network effect propagates. It also enables us to derive a refined upper bound and a matching lower bound (Section 4), making it an order-optimal measure for arbitrary network structures. However, $\mathcal{B}(G;\rho)$ lacks a simplified expression. Although the sum of discounted infinite terms looks like a Neumann series, it is not exactly one because of the Hadamard product. On the other hand, $(1/4)\cdot(\mathbf{I}-\rho^2\tilde{\mathbf{A}})^{-1}\mathbf{b}$ offers a looser bound but offers better interpretability through its connection to a closed-form expression that resembles the inverse in-degree centrality. Because $\tilde{\mathbf{A}}$ is row stochastic, $\boldsymbol{\kappa}(t)$ is bounded by (approximately) the weighted sum of inverse in-degrees $\mathbf{b}$. This implies that as the number of in-neighbors increases, the variance decreases. This is because having more in-neighbors reduces the impact of a neighbor, thus reducing the mutual dependence among adoptions. As time progresses, this upper bound increases, reflecting a discounted contribution of more distant neighbors. As such, although this loose upper bound sacrifices some details, it still offers valuable insights into how stochasticity spreads.

With the bound on the variance of the local network effect, we then move on to bound the nonlinear dynamics. Although the adoption probability $\{\mathbf{q}(t)\}_{t=0}^{\infty}$ lacks a closed-form expression, we expect its transition between consecutive time steps to be similar to the AEO $\mathbf{h}(\cdot)$.

Lemma 3 (Fixed-Point Deviation of Adoption Probability). *Under Assumptions 1 and 2, for any diffusion instance* $(G,\mathbf{v},F_\epsilon(\cdot),\beta)$ *and* $t\geq 1$, *we have*

$$|\mathbf{h}(\mathbf{q}(t-1))-\mathbf{q}(t)|_{\text{ew}}\leq\frac{\rho}{2}[\mathcal{B}(G;\rho)]_{\text{ew}}^{\frac{1}{2}}\leq\frac{\rho}{2}\left[(\mathbf{I}-\rho^2\tilde{\mathbf{A}})^{-1}\mathbf{b}\right]_{\text{ew}}^{\frac{1}{2}}.$$

Based on Lemma 2, Lemma 3 essentially connects the transitions of $\{\mathbf{q}(t)\}_{t=0}^{\infty}$ and $\{\boldsymbol{\mu}(t)\}_{t=0}^{\infty}$, offering a one-step guarantee for them. We then use a "fixed-point sandwich" technique to prove the results in Theorem 1. Specifically, let $\boldsymbol{\delta}:=(\rho/2)\cdot[\mathcal{B}(G;\rho)]_{\text{ew}}^{\frac{1}{2}}$ denote the absolute deviation. We define a lower-bound system $\{\underline{\boldsymbol{\mu}}(t)\}_{t=0}^{\infty}$ and an upper-bound system $\{\bar{\boldsymbol{\mu}}(t)\}_{t=0}^{\infty}$ as follows; for all $i\in V$,

$$\underline{\mu}_i(t)=\begin{cases}q_i(0) & t=0\\ 1-F_\epsilon\left(-v_i-\beta\dfrac{\sum_{j\in\mathcal{N}_i}\underline{\mu}_j(t-1)}{d_i}\right)-\delta_i & t>0,\end{cases}$$

$$\overline{\mu}_i(t)=\begin{cases}q_i(0) & t=0\\ 1-F_\epsilon\left(-v_i-\beta\dfrac{\sum_{j\in\mathcal{N}_i}\overline{\mu}_j(t-1)}{d_i}\right)+\delta_i & t>0.\end{cases}$$

Using auxiliary AEOs, these two systems can be expressed as fixed-point iterations: $\underline{\boldsymbol{\mu}}(t)=\mathbf{h}_{-\boldsymbol{\delta}}(\underline{\boldsymbol{\mu}}(t-1))$ and $\bar{\boldsymbol{\mu}}(t)=\mathbf{h}_{\boldsymbol{\delta}}(\bar{\boldsymbol{\mu}}(t-1))$, with $\underline{\boldsymbol{\mu}}^*$ and $\bar{\boldsymbol{\mu}}^*$ being the respective fixed-point solutions. Note that $\mathbf{h}_{-\boldsymbol{\delta}},\mathbf{h}_{\boldsymbol{\delta}}\in\mathcal{H}$. Then, we leverage these two fixed-point iterations to sandwich both $\{\mathbf{q}(t)\}_{t=0}^{\infty}$ and $\{\boldsymbol{\mu}(t)\}_{t=0}^{\infty}$ and show that both $\underline{\boldsymbol{\mu}}(t)\leq\mathbf{q}(t)\leq\bar{\boldsymbol{\mu}}(t)$ and $\underline{\boldsymbol{\mu}}(t)\leq\boldsymbol{\mu}(t)\leq\bar{\boldsymbol{\mu}}(t)$ hold. Further, one can argue that $\mathbf{0}\leq\bar{\boldsymbol{\mu}}(t)-\underline{\boldsymbol{\mu}}(t)\leq 2(\mathbf{I}-\rho\tilde{\mathbf{A}})^{-1}\boldsymbol{\delta}$. Therefore, $|\mathbf{q}(t)-\boldsymbol{\mu}(t)|_{\text{ew}}\leq 2(\mathbf{I}-\rho\tilde{\mathbf{A}})^{-1}\boldsymbol{\delta}$. Sending t to infinity, the result of Theorem 1 holds. For details, refer to Online Appendix B.3.

We highlight that although for clarity in Theorem 1, we frame our theoretical result in terms of the limiting behavior (i.e., $\boldsymbol{\mu}^*$ and $\mathbf{q}^*$), our arguments here show that these bounds apply to the entire diffusion trajectory. In particular, the fixed-point iteration $\boldsymbol{\mu}(t)$ also approximates the transient adoption probability $\mathbf{q}(t)$ even when t is small, and the approximation error is naturally bounded by the gap between their limiting behaviors, thereby extending the result to the full diffusion process.

4. Order-Optimal Error Bounds for the FPA Scheme

In this section, we delve deeper into the error bound of the FPA scheme. By introducing an additional mild assumption, we derive a tighter upper bound and a matching lower bound of the same order, thus closing the gap in our analysis. For the subsequent analysis, we proceed under Assumption 3.

Assumption 3 (Stronger Smoothness Condition). *The random noise* $\epsilon_i(t)$ *has a differentiable probability density function* $f_\epsilon(\cdot)$ *with its derivative upper bounded by* $|f'_\epsilon(\cdot)|\leq L_f$.

This assumption mainly requires the smoothness of the PDF $f_\epsilon(\cdot)$. It is worth noting that this assumption is fairly mild given that many commonly used distributions inherently exhibit high degrees of differentiability, including but not limited to the normal and logistic distributions.

4.1. Improved Upper Bounds

Recall that Theorem 1 establishes an upper bound for the approximation error at the order of $[\mathcal{B}(G;\rho)]_{\text{ew}}^{\frac{1}{2}}$. Under Assumption 3, Theorem 2 refines this upper bound, improving it to a tighter order of $\mathcal{B}(G;\rho)$. Define the constant $\tilde{C}:=L_f\beta^2/[4(1-\rho)^2]$, which increases with both ρ and L_f.

Theorem 2 (Improved Entry-Wise Error Bound of the FPA Scheme). *Under Assumptions 1, 2, and 3, for any diffusion instance* $(G,\mathbf{v},F_\epsilon(\cdot),\beta)$, *we have*

$$|\mathbf{q}^*-\boldsymbol{\mu}^*|_{\text{ew}}\leq\frac{L_f\beta^2}{4}(\mathbf{I}-\rho\tilde{\mathbf{A}})^{-1}\mathcal{B}(G;\rho)\leq\tilde{C}\cdot\mathcal{C}(G;\rho)$$

$$\textit{and}\quad\|\mathbf{q}^*-\boldsymbol{\mu}^*\|_\infty\leq\tilde{C}\cdot\frac{1}{d_{\min}}.\tag{14}$$

We observe that both $\mathcal{B}(G;\rho)$ and the inverse in-degree centrality $\mathcal{C}(G;\rho)$ remain crucial in the improved bound. Unlike Theorem 1, the dependency of the bound on these measures is now linear. In light of the refined bound (14), we also sharpen the scaled ℓ_1-norm of the approximation error.

Corollary 3 (Improved Scaled ℓ_1-Norm Bound). *Under Assumptions 1, 2, and 3, for any diffusion instance* $(G,\mathbf{v},F_\epsilon(\cdot),\beta)$ *with* $r(G)<1/\rho$, *we have*

$$\frac{1}{n}\|\mathbf{q}^*-\boldsymbol{\mu}^*\|_1 \le \frac{(1-\rho)\tilde{C}}{1-\rho r(G)}\cdot\mathcal{D}(G). \tag{15}$$

The proof of Theorem 2 largely parallels that of Theorem 1, with the key difference being the use of a second-order Taylor expansion of F_ϵ, to bound $|\mathbf{h}(\mathbf{q}(t-1))-\mathbf{q}(t)|_{\text{ew}}$ (see Lemma EC.6 in Online Appendix C).

4.2. A Matching Lower Bound

Our results in Section 4.1 highlight a linear dependence of the FPA scheme's error upper bounds on network measures, such as $\mathcal{B}(G;\rho)$, $\mathcal{C}(G;\rho)$, $1/d_{\min}$, and $\mathcal{D}(G)$. We now establish matching lower bounds of the same order, suggesting that these bounds are order optimal.

Theorem 3 (Entry-Wise Lower Bound of the Error). *For any given network G, there exists a diffusion instance* $(G,\mathbf{v},F_\epsilon(\cdot),\beta)$ *satisfying Assumptions 1, 2, and 3 such that*

$$|\mathbf{q}^*-\boldsymbol{\mu}^*|_{\text{ew}} \ge \hat{C}(\mathbf{I}-\rho_\ell\tilde{\mathbf{A}})^{-1}\mathcal{B}(G;\rho_\ell), \tag{16}$$

where constants $\hat{C},\rho_\ell$ *depend on* $\mathbf{v}$, $F_\epsilon(\cdot)$, *and* β.

Notably, the constants $\tilde{C}$ and ρ in (14) depend on the distribution of random error, particularly the upper continuity bounds. The constants in the matching lower bound (i.e., $\hat{C}$ and ρ_ℓ) instead depend on the lower continuity bounds. Specifically, to establish the lower bound (16), for any $\xi>0$, we construct a diffusion instance, where (i) each agent $i\in V$ has the same intrinsic value $v=-\beta-\xi$ and (ii) the random noise follows $\epsilon_i(t)\sim$ Logistic$(0,1)$. A key consequence is that for any node $i\in V$, $v_i+\beta\sum_{j\in\mathcal{N}_i}Y_j(t-1)/d_i\in[v,v+\beta]=[-\beta-\xi,-\xi]$, so we can set ℓ and ℓ_f as arbitrary values satisfying $|F_\epsilon(x)-F_\epsilon(y)|\ge\ell|x-y|$ and $|f'_\epsilon(x)|\ge\ell_f$, respectively, for all x, y in $[-\beta-\xi,-\xi]$. These bounds allow us to reverse the chain of inequalities in the analysis of upper bounds, eventually leading to the establishment of (16), with $\hat{C}=\ell_f\beta^2[1-F_\epsilon(\xi+\beta)]F_\epsilon(\xi)/2$ and $\rho_\ell=\beta\ell$. Furthermore, when the underlying graph is a regular graph, we can also establish the matching lower bounds with regard to other proposed network metrics (i.e., $\mathcal{C}(G;\rho)$, $1/d_{\min}$, and $\mathcal{D}(G)$). We comment that the constants in the lower and upper bounds do not match, and we relegate the analysis regarding the (sub-)optimality of constants to future research.

5. Numerical Experiments

In this section, we present numerical studies to validate our FPA scheme in different scenarios, including the 10-node example (i.e., Figure 1(a)), large-scale random networks, and real-world networks. To help readers better understand our numerical results, we provide the code for our analyses in a GitHub repository (https://github.com/YunduanLin/Nonprogressive_Diffusion). Our results reveal several key insights. First, the FPA scheme consistently achieves superior performance, even for small and sparse networks. Second, the inverse in-degree centrality $\mathcal{C}(G;\rho)$ and the inverse in-degree intensity $\mathcal{D}(G)$ emerge as strong indicators of FPA performance. Third, in terms of computational efficiency, the FPA scheme significantly outperforms alternative simulation approaches. Lastly, a heuristic mixture scheme, which combines the FPA scheme with resampling, further improves approximation performance with only a slight increase in computational overhead.

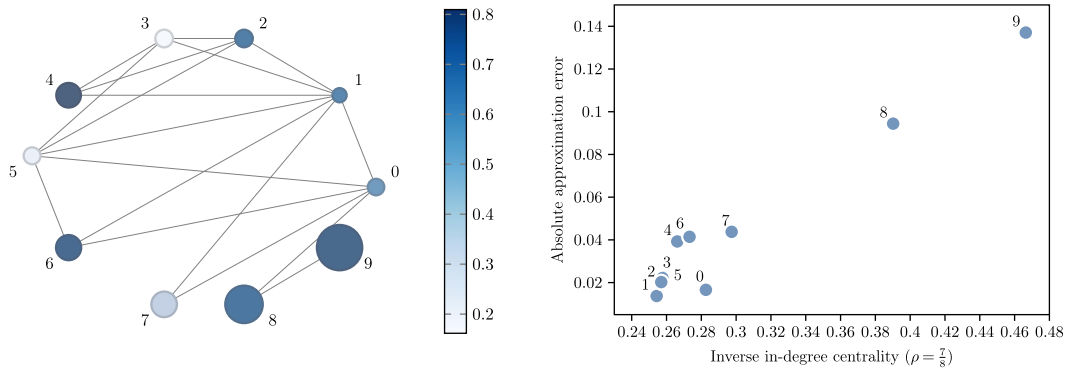
A key challenge in evaluating FPA performance is that the limiting adoption probability $\mathbf{q}^*$ is generally unknown as computing it requires solving the stationary distribution of a large-scale MC (see Section 2), which is typically infeasible.[4] As a practical workaround, we estimate the ground-truth $\mathbf{q}^*$ (see Online Appendix D.3) via agent-based simulation over a long time horizon. Unless otherwise specified, we adopt the following setups; $\mathbf{q}^*$ is estimated by ABS, whereas $\boldsymbol{\mu}^*$ is obtained through fixed-point iteration, with an initial value $\boldsymbol{\mu}(0)=\mathbf{0}$ and a convergence criterion of 10^{-5}. We set $\epsilon_i(t)\overset{\text{i.i.d.}}{\sim}$ Logistic$(0,1)$ and $\beta=3.5$, so $\rho=0.875$. This relatively large ρ allows us to test FPA in near-worst-case scenarios. To measure the performance of a specific diffusion instance, we use the mean absolute percentage error across all agents defined as $\text{MAPE}=(1/n)\cdot\sum_{i\in V}|\mu_i^*-q_i^*|/q_i^*\times100\%$. This self-normalized measure facilitates interpretation and comparison across different instances.

5.1. Revisiting the Motivating Example

In this subsection, we revisit the 10-node example introduced in Section 3.1. We mainly focus on the role of our centrality measure as a node-level metric. We also include a thorough examination of Assumption 2 based on this example; see Online Appendix D.1.1. For experiments in this subsection, $\mathbf{q}^*$ is achieved by solving the stationary distribution of the MC because of the manageable size.

Our theorems link the upper bound of the approximation error to inverse in-degree centrality $\mathcal{C}(G;\rho)$. In Figure 2, we juxtapose the approximation error $|q_i^*-\mu_i^*|$ with $\mathcal{C}_i(G,\rho)$ for all 10 nodes. The result reveals a clear positive correlation, emphasizing the role of inverse in-degree centrality as a sharp node-level indicator to evaluate FPA performance.

Figure 2. (Color online) Analysis of the 10-Node Example Instance



Notes. (Left panel) Reproduction of Figure 1(a) for reference. (Right panel) Illustration of the relationship between the absolute approximation error and inverse in-degree centrality.

5.2. Random Networks

In this section, we evaluate FPA performance across various well-studied random networks, focusing on Erdős–Rényi networks of different sizes and densities and power-law networks with different exponents and degree correlations (see Online Appendix D.4) (Erdős and Rényi 1960). To ensure robustness, we conduct 50 repetitions for each combination of random network parameters. In the following, we consider a sequence of directed Erdős–Rényi networks, each denoted by $G(n, p(n))$, where n represents the network size and $p(n)$ represents the network density—the probability that an edge between two nodes exists (Erdős and Rényi 1960). The presence of edges is assumed to be independent. Our analysis examines both the sensitivity of FPA performance to network structure and the computational efficiency of the FPA scheme.

5.2.1. FPA Performance with Regard to Network Structure.

We assess FPA performance by varying the size and density of Erdős–Rényi networks, respectively (Erdős and Rényi 1960).

i. *Network size*. We vary network size n from 20 to 10,000 and select network density $p(n) \in \{1/n^{1.1}, 1/n, (\log n)^2/n, 0.1\}$. These values are chosen based on the critical ranges of $p(n)$ identified in random graph theory. Further details on the properties of Erdős–Rényi networks are provided in Online Appendix D.5 (Erdős and Rényi 1960; see also Huang et al. 2022). Although our theoretical results assume $d_{\min} > 0$, Erdős–Rényi networks may contain stand-alone nodes with zero in-degree (Erdős and Rényi 1960). These nodes, which do not receive influence but can exert influence (i.e., their out-degree can be positive), are perfectly approximated by the FPA scheme. To align with theory, we exclude stand-alone nodes from calculation in this subsection, but we include them elsewhere to reflect a more realistic evaluation of real-world network.

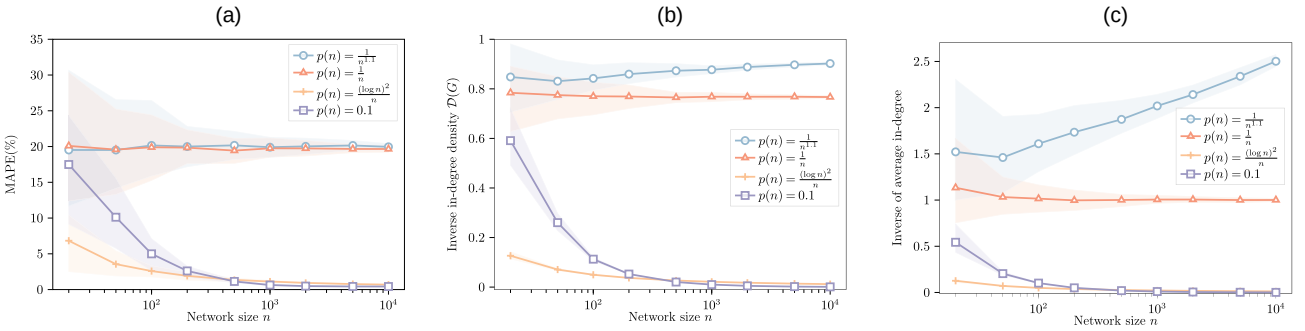
In Figure 3(a), we examine how MAPE varies with network size n. For all tested cases, the MAPE either decreases or remains stable as n grows, with a more pronounced decrease when $p(n)$ is large. Panels (b) and (c) of Figure 3 present two metrics—the inverse in-degree density $\mathcal{D}(G)$ and the inverse of average in-degree. As n increases, we find that $\mathcal{D}(G)$ trends similarly to MAPE, indicating that it is a strong indicator of FPA performance. In contrast, the average in-degree, which traditionally indicates density, shows weaker explanatory power for this trend.

ii. *Network density*. We fix a medium network size of $n = 1,000$ and focus on various densities $p(n) \in \{1/n^{1.3}, 1/n^{1.1}, 1/n, \sqrt{\log n}/n, \log n/n, (\log n)^2/n, 0.1\}$. In Figure 4(a), the MAPE decreases as the network gets denser, reflecting a trend that aligns with Figure 3(a) and our theoretical results.

Recall that we set ρ close to one to demonstrate the near-worst-case performance. Even under such a setting, the FPA scheme performs exceptionally well. Panels (b) and (c) of Figure 4 further show two network metrics: $\mathcal{D}(G)$ and the inverse of average in-degree, respectively. The trends largely mirror those seen in panels (b) and (c) of Figure 3, again confirming that $\mathcal{D}(G)$ is an informative indicator of FPA performance.

We observe that the improved FPA performance in dense networks can also be partially attributed to the largest out-degree-to-in-degree ratio $r(G)$. As highlighted in Corollary 2, the upper bound for the approximation error increases as $r(G)$ increases. In dense Erdős–Rényi networks, both in-degrees and out-degrees of nodes cluster around the mean value, promoting both density and balance of the network (Erdős and Rényi 1960). To further explore the role of network imbalance, we extensively analyze power-law networks with varying in-degree/out-degree correlations (see Online Appendix D.4).

Figure 3. (Color online) FPA Performance on Erdős–Rényi Networks of Different Network Sizes



Notes. All horizontal axes are on the log scale. Shaded areas represent the 95% confidence intervals. (a) MAPE against n. (b) Calligraphic D(G) against n. (c) Inverse of average in-degree against n.

5.2.2. Computational Efficiency of the FPA Scheme.

We evaluate the computational efficiency by comparing the central processing unit (CPU) time required to compute $\boldsymbol{\mu}^*$ with that needed to estimate $\mathbf{q}^*$ using two simulation methods. The first is the naïve ABS method, which simulates adoption dynamics based on the agent behavior model defined in (1) and estimates $\mathbf{q}^*$ by averaging adoption states after the warm-up period as described in the first equation of (3). To ensure a fair comparison, we report the simulation time for the naïve ABS method once its real-time MAPE falls below the value achieved by the FPA scheme. Although the naïve ABS method is a natural approach for simulating limiting adoption probability, it can be time consuming and does not fully capture the performance of advanced simulation techniques. Thus, we also include an accelerated agent-based simulation (A-ABS) method, which improves simulation efficiency by leveraging the analytical form of the behavior model. A-ABS estimates the adoption probability by averaging conditional probabilities as described in the second equation of (3). More details on our simulation benchmarks are provided in Online Appendix D.3.

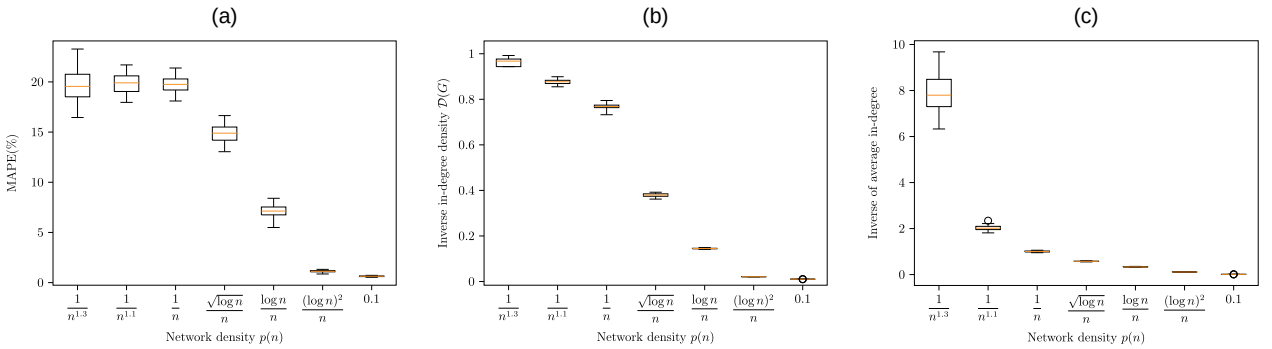
We fix the network density at $p(n) = 0.1$ and vary n from 20 to 10,000 to compare the run time. As shown in Table 1, the run time for all methods increases with n, but the FPA scheme consistently outperforms both simulation methods by a substantial margin, with the performance gap widening for larger networks. When $n = 10,000$, the FPA scheme completes in 2.3 seconds compared with 40 minutes for the naïve ABS method and 6.5 minutes for the A-ABS method to achieve the same performance, making it roughly 1,027 times and 170 times more efficient, respectively.

In conclusion, the FPA scheme offers considerable advantages in computational efficiency across all of the tested scenarios without a significant compromise in accuracy.

5.3. Real-World Networks

Our numerical experiments in Section 5.2 focus on random networks, which may not capture real-world phenomena (e.g., see Jackson 2010). To evaluate the FPA scheme in more realistic settings, we test it on real-world networks from the Network Repository (Rossi and Ahmed 2015). Specifically, we select five social friendship networks extracted from Facebook, where nodes represent individuals and edges denote friendship ties. A summary of these networks and their results is provided in Table 2. For raw data and

Figure 4. (Color online) FPA Performance on Erdős–Rényi Networks of Different Network Densities



Notes. (a) MAPE against $p(n)$. (b) Calligraphic D(G) against $p(n)$. (c) Inverse of average in-degree against $p(n)$.

Table 1. The CPU Time Required for Simulation Methods and the FPA Scheme

Network size n	20	50	100	200	500	1,000	2,000	5,000	10,000
Naïve ABS time (s)	0.168	0.258	0.530	2.415	18.979	97.452	286.718	1,315.301	2,366.879
A-ABS time (s)	0.152	0.285	0.512	0.996	2.380	5.456	15.083	93.235	391.357
FPA time (s)	0.002	0.004	0.007	0.018	0.044	0.102	0.230	1.027	2.304

Note. s, seconds.

additional statistics, readers can refer to the Network Repository website.[5]

We emphasize three observations from these experiments. First, the FPA scheme performs exceptionally well, achieving a maximum MAPE of just 3.48% across all tested networks, which indicates both accuracy and reliability for real-world applications. Second, the FPA scheme significantly outperforms the naïve ABS method in terms of computational time, with a speedup factor ranging from 70 to 230. Even compared with the A-ABS method, the FPA scheme is 23–30 times faster, showcasing its efficiency and scalability. Third, among various metrics, the inverse in-degree density $\mathcal{D}(G)$ stands out as the most reliable indicator of FPA performance measured by MAPE.

Figure 5 presents a comprehensive overview of the relationship between FPA performance and the inverse in-degree density across different networks. Figure 5 reveals a clear positive correlation between MAPE and $\mathcal{D}(G)$, reinforcing that $\mathcal{D}(G)$ is not just an upper bound of FPA performance but also, a reliable performance metric. Interestingly, real-world networks generally exhibit lower $\mathcal{D}(G)$ values and MAPE for the FPA scheme compared with most random networks. This suggests that random networks may not fully capture the characteristics of social networks in our setting, further emphasizing the practical relevance of our FPA scheme for arbitrary network structures.

5.4. A Heuristic Mixture of the FPA and Simulation Scheme

Although the FPA scheme improves time efficiency in network diffusion characterization, it sacrifices precision, particularly for low in-degree agents, as shown both theoretically and empirically in Online Appendix D.5. This limitation stems from the nonlinear transformation during the diffusion process. Even if the FPA scheme accurately approximates the adoption probabilities of an agent's in-neighbors, it may not be accurate for the agent itself because of the deterministic nature. To mitigate this, we propose a heuristic mixture scheme that combines the FPA scheme with resampling, improving accuracy with a slight increase in computational overhead.

In this mixture scheme, we first derive the FPA solution $\boldsymbol{\mu}^*$. Next, we resample the adoption states of agent i as independent Bernoulli variables $\tilde{Y}_i$ with probability μ_i^*. The adoption probability for agent i can then be refined through simulation by evaluating the following expected value:

$$\tilde{q}_i = \mathbb{E}\left[1 - F_\epsilon\left(-v_i - \beta\frac{\sum_{j\in\mathcal{N}_i}\tilde{Y}_j}{d_i}\right)\right]. \tag{17}$$

In terms of precision, the mixture scheme notably improves precision for low in-degree agents with high-degree neighbors; the FPA scheme accurately approximates the neighbors' adoption states, whereas resampling adds necessary stochasticity for the focal agent. In terms of efficiency, the mixture scheme remains scalable as resampling can be parallelized without a full simulation. Moreover, it can be applied selectively to low-degree agents, avoiding unnecessary computation for all agents. This balance ensures improved accuracy while maintaining minimal computational overhead.

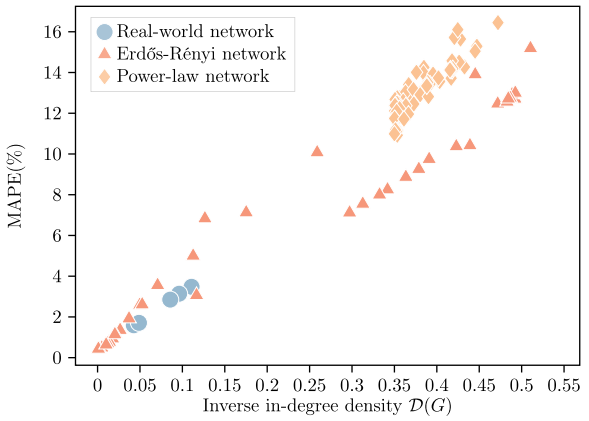
When implementing this mixture scheme, two key factors must be considered: (1) which agents to resample and (2) the sample size for estimating $\tilde{\mathbf{q}}$ as shown in (17). To illustrate its performance, we evaluate the MAPE of the mixture scheme on two real-world networks, Caltech36 and Amherst41, as introduced in Section 5.3. We vary the in-degree cutoff from 0 to 200, applying the mixture scheme only to agents

Table 2. Experiment Results for Real-World Networks

							Computation time (s)		
Instance	n	$d_{\min}$	$d_{\max}$	$\bar{d}$	$\mathcal{D}(G)$	MAPE (%)	Naïve ABS	A-ABS	FPA
Caltech36	770	1	248	43.2623	0.1108	3.48	4.5335	1.5075	0.0636
Reed98	963	1	313	39.0696	0.0962	3.14	5.6228	1.9044	0.0623
Haverford76	1,447	1	375	82.3621	0.0427	1.59	23.2347	2.6996	0.1009
Simmons81	1,519	1	300	43.4338	0.0857	2.85	11.9185	3.4207	0.1426
Amherst41	2,236	1	467	81.3542	0.0488	1.71	35.6749	4.7757	0.1846

Note. s, seconds.

Figure 5. (Color online) FPA Performance Against the Inverse In-Degree Density for Different Networks



Note. Each point for the random networks represents the average value of the same parameter pair across all repetitions.

below each threshold, and we consider sample sizes of {100, 1,000}. Notably, when the cutoff degree is set to zero, the mixture scheme reduces to the FPA scheme alone.

In Figure 6, we present the performance of the mixture scheme under various conditions. The results show that the mixture scheme consistently outperforms the FPA scheme across all scenarios, with the most notable improvements observed from low in-degree agents. This is reflected in the sharp MAPE drop at low cutoff values. Moreover, larger sample sizes, which provide more accurate estimations of (17), lead to better performance and more flexible cutoff selection. With a sample size of 100, MAPE initially decreases as the cutoff in-degree increases, minimizing around a cutoff of 20, and then, it increases in both instances. This suggests that coarse estimations of $\tilde{\mathbf{q}}$ work well for low in-degree agents but can negatively affect high in-degree agents who are well approximated by the FPA scheme, emphasizing the importance of cutoff selection. When the sample size is increased to 1,000, which provides a more accurate evaluation of $\tilde{\mathbf{q}}$, MAPE decreases almost monotonically as the cutoff in-degree increases, reducing to around 1% for both instances.

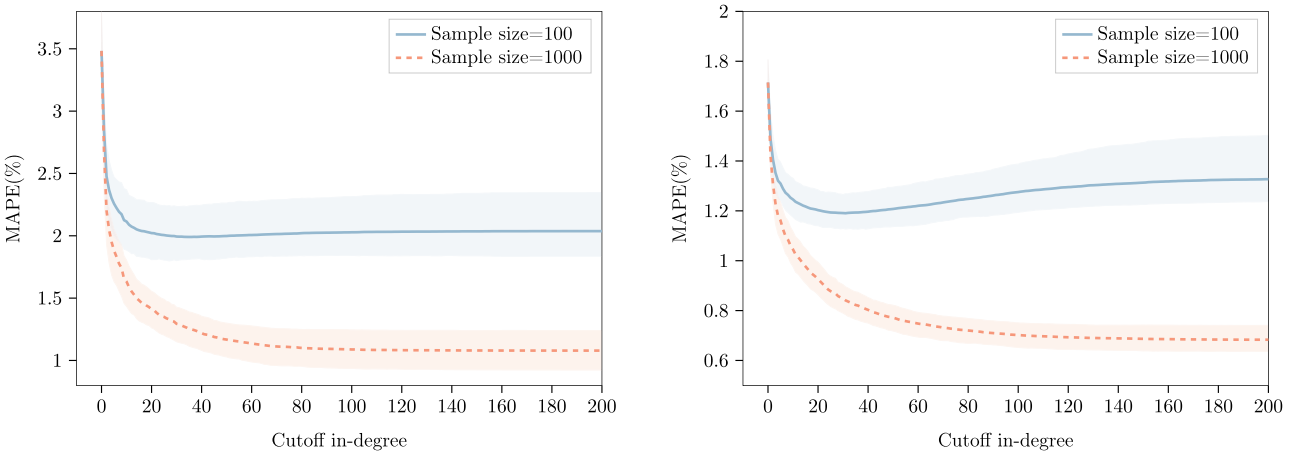
Notably, the warm-up period for ABS methods is set to 1,000. By parallelizing computations and targeting only low in-degree agents, the mixture scheme can be significantly more efficient than the warm-up period of ABS methods. However, the MAPE of the mixture scheme does not reach zero as it relies on independent resampling. Although the actual adoption states of different agents are positively correlated, we cannot capture the full joint distribution of adoption states.

In Figure 7, we present a comprehensive comparison of the mixture scheme across various synthetic and real-world networks alongside the FPA scheme. The mixture scheme consistently shows significant improvements, particularly in cases where the FPA scheme performs poorly. On average, using a cutoff in-degree of 20 and a sample size of 100, the mixture scheme reduces the MAPE by 47.08%, whereas a cutoff in-degree of 200 and a sample size of 1,000 achieve a reduction of 58.94%. In conclusion, the resampling approach effectively complements the FPA scheme. In practice, the cutoff in-degree and sample size can be adjusted to balance precision and efficiency.

6. Applications of the Fixed-Point Approximation Scheme

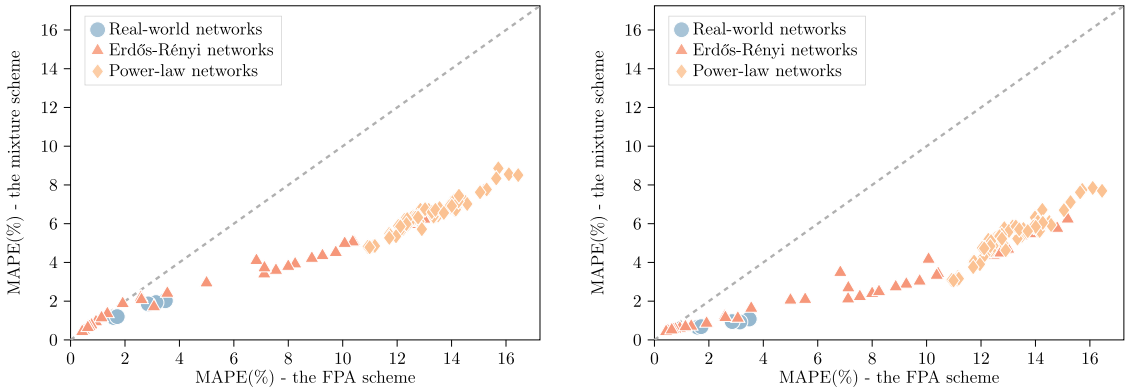
The FPA scheme can be applied to many classical operational decision problems involving network diffusion. In this section, we consider two notable

Figure 6. (Color online) Performance of the Mixture Scheme Under Different Conditions



Notes. (Left panel) Caltech36. (Right panel) Amherst41. Shaded areas represent the 95% confidence intervals.

Figure 7. (Color online) Performance Comparison Between the Mixture Scheme and the FPA Scheme



Notes. (Left panel) Sample size of 100. (Right panel) Sample size of 1,000. The gray dashed line represents $y = x$, indicating equal performance of the two schemes.

applications: the influence maximization problem in network analysis and optimal pricing in revenue management. Hereafter, we confine our analysis of the optimization problems to a given instance $(G, \mathbf{v}, F_\epsilon(\cdot), \beta)$. As the approximation error of the FPA scheme does not translate into the optimality gap of the optimization problem directly, we include a regret analysis in Online Appendix E.1.

6.1. Influence Maximization

In the IM problem, the goal is to select up to K seed users to initially adopt the service so as to maximize long-term adoptions of the entire network. For example, a service provider may target key influencers to trigger broader diffusion. We assume that seed adoptions are irreversible, unlike typical nonprogressive settings, for two reasons; (i) it reflects the lasting commitment of seed users in practice, and (ii) as shown in Proposition 1, altering only initial states does not affect the long-term equilibrium. By making seed adoptions irreversible, we effectively modify limiting behavior. This can also be seen as assigning sufficiently high intrinsic values to seed users to guarantee adoptions.

Given diffusion instance $(G, \mathbf{v}, F_\epsilon(\cdot), \beta)$, the original IM problem can be formulated as

$$\underset{S\subseteq V:|S|=K}{\text{maximize}} \quad \sum_{i\in V} q_i^* = \lim_{t\to\infty} \sum_{i\in V} \mathbb{E}[Y_i(t)] \tag{18a}$$

$$\text{subject to} \quad Y_i(t) = \begin{cases} 1 & \forall i \in S, t \geq 1, \\ \mathbb{1}\left\{ v_i + \beta \dfrac{\sum_{j\in\mathcal{N}_i} Y_i(t-1)}{d_i} + \epsilon_i(t) \geq 0 \right\} & \forall i \in V \setminus S, t \geq 1. \end{cases} \tag{18b}$$

Objective (18a) is the limiting total expected adoptions, and (18b) describes the MC that determines $\mathbf{q}^*$ with initialization $\mathbf{Y}(0) = \mathbf{1}$ and $Y_i(t) = 1$ for all $i \in S$ and $t \geq 1$.

Employing the FPA scheme, the approximate IM problem can be formalized as follows:

$$\underset{\boldsymbol{\mu}, S\subseteq V:|S|=K}{\text{maximize}} \quad \sum_{i\in V} \mu_i \tag{19a}$$

$$\text{subject to} \quad \mu_i = 1, \qquad \forall i \in S \tag{19b}$$

$$\mu_i = 1 - F_\epsilon\left(-v_i - \beta \frac{\sum_{j\in\mathcal{N}_i} \mu_j}{d_i} \right), \quad \forall i \in V \setminus S. \tag{19c}$$

For ease of formulation, we use $\boldsymbol{\mu}$ as an explicit decision variable, and a set of equality constraints specifies the FPA scheme, which uniquely determines $\boldsymbol{\mu}^*(S)$ for any given $S \subseteq V$.

Although (19a) provides a good approximation to (18), solving it remains challenging. Under a mild condition stated below, we prove that $g(\boldsymbol{\mu}^*(S), S)$ is submodular with regard to the seed set S.

Assumption 4 (Restricted Convexity of the CDF). *The random noise CDF $F_\epsilon(\cdot)$ is convex on the domain* $[-\max_{i\in V} v_i - \beta, -\min_{i\in V} v_i]$.

Assumption 4 covers a wide range of cases commonly studied, including the nonprogressive LT model as a special case of our model. Further discussion and additional examples supporting this assumption can be found in Online Appendix E.2.

Theorem 4 (Submodularity of the Approximate IM Problem). *Under Assumptions 1, 2, and 4, $g(\boldsymbol{\mu}^*(S), S)$ is a submodular set function with regard to seed set S.*

Theorem 4 enables the use of a well-known greedy algorithm (e.g., see Nemhauser et al. 1978), which iteratively adds the node with the largest marginal gain in total adoptions, to solve the approximate IM problem. For instances that satisfy Assumption 4, the greedy algorithm provides a $(1 - 1/e)$-approximation solution to Problem (19a). Together with regret analysis

(Proposition EC.1 in Online Appendix E.1), the simple greedy approach provides a high-quality solution to the original IM Problem (18).

In summary, our approximate IM formulation offers several advantages. It allows us to establish a clear condition (i.e., Assumption 4), ensuring submodularity, which is hard to verify for the original problem. Furthermore, it enables a more efficient greedy algorithm as the FPA scheme accelerates the computation of adoption probabilities needed $\mathcal{O}(nK)$ times. To validate our findings, we conduct extensive numerical experiments, with details provided in Online Appendix E.3. Our results show that regardless of Assumption 4, the greedy algorithm for the approximate IM problem achieves near-optimal solutions. Moreover, it outperforms many heuristics and performs nearly as well as the greedy solution for the original problem while running much faster.

6.2. Optimal Pricing on a Social Network

Network effects play an important role in shaping customer preferences for products or services, which have driven a growing body of literature that integrates network effects into revenue management problems (Du et al. 2016, 2018; Wang and Wang 2017; Chen and Chen 2021; Chen and Shi 2023; Gopalakrishnan et al. 2023). Our model naturally connects to this literature, which often employs axiomatic or game-theoretic frameworks to approximate the limiting behavior in dynamic, stochastic environments. In this section, we first follow the literature by discussing the general static pricing problem under limiting behaviors. We then extend this framework to dynamic pricing strategies for transient periods, assuming that the platform can exercise perfect price discrimination. This is made possible by the theoretical results in Section 3.3; the FPA scheme captures not only the limiting behavior but also, the full transient diffusion trajectory.

6.2.1. Static Pricing Problem Under Limiting Behavior.

We consider a firm using pricing as an operational lever to steer consumers' adoption decisions. The adoption utility for user i at time t is given by $u_i(t) = v_i - \gamma p_i + \beta \sum_{j\in\mathcal{N}_i} Y_j(t-1)/d_i + \epsilon_i(t)$, where p_i is the price offered to user i and γ denotes the price sensitivity. We allow for different prices to be offered to different consumers as many platforms are capable of targeted price discrimination. Suppose the platform can offer up to m distinct prices, represented by $\mathbf{p} \in \mathbb{R}^m$, using a known transformation matrix $\mathbf{W} \in \mathbb{R}^{n\times m}$, where $W_{ik} = 1$ if consumer i is assigned to the kth price and $W_i = 0$ otherwise. When $m = n$ and $\mathbf{W} = \mathbf{I}_n$, customers face idiosyncratic prices. When $m = 1$ and $\mathbf{W} = \mathbf{e}_n$, customers face a homogeneous price. Intermediate segmentations, such as pricing by network connectivity, are also possible. The ultimate objective is to identify an optimal price vector that maximizes total profit.

Given diffusion instance $(G, \mathbf{v}, F_\epsilon(\cdot), \beta)$, the original pricing problem can be formulated as

$$\underset{\mathbf{p}}{\text{maximize}} \quad \sum_{i\in V}\left(\sum_{k=1}^{m} W_{ik}p_k\right)\cdot q_i^* = \sum_{i\in V}\left(\sum_{k=1}^{m} W_{ik}p_k\right) \cdot \lim_{t\to\infty}\mathbb{E}[Y_i(t)] \tag{20a}$$

$$\text{subject to} \quad Y_i(t) = \mathbb{1}\left\{v_i - \gamma\sum_{k=1}^{m} W_{ik}p_k + \beta\frac{\sum_{j\in\mathcal{N}_i} Y_i(t-1)}{d_i} + \epsilon_i(t) \ge 0\right\} \ \forall i \in V, t \ge 1. \tag{20b}$$

Objective (20a) is the total profit, and (20b) describes the MC that determines $\mathbf{q}^*$ with given price $\mathbf{p}$. Employing the FPA scheme, the approximate problem can be formally stated as

$$\underset{\boldsymbol{\mu},\mathbf{p}}{\text{maximize}} \quad \boldsymbol{\mu}^\top \mathbf{W}\mathbf{p} \tag{21a}$$

$$\text{subject to} \quad \mu_i = 1 - F_\epsilon\left(-v_i + \gamma\sum_{k=1}^{m} W_{ik}p_k - \beta\frac{\sum_{j\in\mathcal{N}_i}\mu_j}{d_i}\right), \ \forall i \in V. \tag{21b}$$

We treat $\boldsymbol{\mu}$ as an explicit decision variable and use (21b) to link $\mathbf{p}$ and $\boldsymbol{\mu}$. The approximate Problem (21a) is generally nonconvex because of Constraint (21b), making it challenging to solve.

The distribution of random noise affects both the formulation and hardness of the problem. Hereafter, we focus on the standard logistic distribution (i.e., $\epsilon_i(t) \overset{\text{i.i.d.}}{\sim} \text{Logistic}(0,1)$). This choice aligns the formulation with existing revenue management literature (Li and Huh 2011, Gallego and Wang 2014, Golrezaei et al. 2020, Chen and Shi 2023). A well-established technique for analyzing such problems is to transform them into optimization problems in demand space. Motivated by this approach, we consider the problem in both the adoption probability and price spaces.

6.2.1.1. Profit Maximization in the Adoption Probability Space. When considering the adoption probability space, the pricing problem becomes less challenging for certain cases. Under perfect price discrimination ($m = n$, $\mathbf{W} = \mathbf{I}_n$) where the platform can provide an idiosyncratic price/subsidy to each consumer and without additional constraints, the problem can be expressed as follows:[6]

$$\underset{\boldsymbol{\mu},\mathbf{p}}{\text{maximize}} \quad \boldsymbol{\mu}^\top \mathbf{p} \tag{22a}$$

$$\text{subject to} \quad \mu_i = 1 - \frac{1}{1 + \exp\{v_i + \beta\sum_{j\in\mathcal{N}_i}\mu_j/d_i - \gamma p_i\}}, \ \forall i \in V. \tag{22b}$$

Cancelling out $\mathbf{p}$, we can reformulate the problem in the adoption probability space as

$$\underset{\boldsymbol{\mu}}{\text{maximize}} \quad \sum_{i\in V}\frac{1}{\gamma}\left(v_i+\beta\sum_{j\in\mathcal{N}_i}\frac{\mu_j}{d_i}+\ln\frac{1-\mu_i}{\mu_i}\right)\mu_i \tag{23a}$$

$$\text{subject to} \quad 0\le\mu_i\le 1,\ \forall i\in V. \tag{23b}$$

When $\beta=0$, the network effect term has no impact, and the problem reduces to the classical pricing problem with a concave objective. We show that this property is preserved when β is small.

Theorem 5 (Concavity of the Static Pricing Objective). *The objective function* (23a) *of the static pricing Problem* (23) *is concave in* $\boldsymbol{\mu}$ *when* $0\le\beta\le 6.25/\lambda_{\max}(\tilde{\mathbf{A}}+\tilde{\mathbf{A}}^\top)$. *When* $\tilde{\mathbf{A}}$ *is a symmetric matrix, the condition reduces to* $0\le\beta\le 3.375$.

Theorem 5 states that when β is appropriately upper bounded, Problem (23) becomes a convex optimization problem, and the optimal solution $\boldsymbol{\mu}^*$ can be achieved by standard optimization techniques (i.e., gradient methods). Once $\boldsymbol{\mu}^*$ is obtained, the optimal prices can be recovered. Furthermore, we remark that both Theorem 5 and Assumption 2 require the network effect intensity to be relatively small.

6.2.1.2. Profit Maximization in the Price Space.

In a more general setting where perfect price discrimination is not feasible, the pricing problem cannot be reformulated in the adoption probability space. Thus, we need to study profit maximization directly in the price space. Particularly, we represent the adoption probability as an implicit function of price, $\boldsymbol{\mu}(\mathbf{p})$, and write the profit function as $\Pi(\mathbf{p})$. We can then derive the gradient of $\Pi(\mathbf{p})$ as follows:

$$\frac{d\Pi(\mathbf{p})}{d\mathbf{p}}=\frac{d\boldsymbol{\mu}(\mathbf{p})}{d\mathbf{p}}\cdot\mathbf{W}\cdot\mathbf{p}+\mathbf{W}^\top\cdot\boldsymbol{\mu}(\mathbf{p}), \tag{24}$$

where the gradient of $\boldsymbol{\mu}(\mathbf{p})$ is not explicitly given. To obtain this, we apply the implicit function theorem to (22b) (i.e., $\boldsymbol{\mu}(\mathbf{p})=\mathbf{h}(\mathbf{p},\boldsymbol{\mu}(\mathbf{p}))$) (see Online Appendix E.4 for details) and rewrite (24) as

$$\frac{d\Pi(\mathbf{p})}{d\mathbf{p}}=\frac{\partial\mathbf{h}(\mathbf{p},\boldsymbol{\mu}(\mathbf{p}))}{\partial\mathbf{p}}\cdot\left(\mathbf{I}-\frac{\partial\mathbf{h}(\mathbf{p},\boldsymbol{\mu}(\mathbf{p}))}{\partial\boldsymbol{\mu}(\mathbf{p})}\right)^{-1}\cdot\mathbf{W}\cdot\mathbf{p}+\mathbf{W}^\top\cdot\boldsymbol{\mu}(\mathbf{p}). \tag{25}$$

We can now apply standard gradient descent techniques to find the near-optimal solution. As a final remark, the profit maximization in the price space as well as the gradient-based approach also holds for other noise distributions or with more complicated price constraints (e.g., the box constraints).

6.2.2. Dynamic Pricing Problem Under Transient Behavior.

Solving the general dynamic pricing problem is challenging, but the special case of perfect price discrimination can also be addressed through reformulation, similar to the static Problem (23). We can then formulate the (approximate) dynamic pricing problem over T time periods as follows:

$$\underset{\boldsymbol{\mu},\mathbf{p}}{\text{maximize}} \sum_{t=1}^{T}\boldsymbol{\mu}(t)^\top\mathbf{p}(t)$$

$$\text{subject to } \mu_i(t)=1-\frac{1}{1+\exp\{v_i+\beta\sum_{j\in\mathcal{N}_i}\mu_j(t-1)/d_i-\gamma p_i(t)\}},\quad \forall i\in V,t\ge 1,$$

where we initialize the adoption probability as $\boldsymbol{\mu}(0)=\mathbf{0}$. Similar to the static problem, we can eliminate $\mathbf{p}$ and reformulate the dynamic problem in the adoption probability space as

$$\underset{\boldsymbol{\mu}}{\text{maximize}} \sum_{t=1}^{T}\sum_{i\in V}\frac{1}{\gamma}\left(v_i+\beta\sum_{j\in\mathcal{N}_i}\frac{\mu_j(t-1)}{d_i}+\ln\frac{1-\mu_i(t)}{\mu_i(t)}\right)\mu_i(t) \tag{26a}$$

$$\text{subject to} \quad 0\le\mu_i(t)\le 1,\ \forall i\in V,t\ge 1. \tag{26b}$$

In Theorem 6 below, we show that the concavity of the objective function is preserved under the same conditions for the dynamic pricing problem as the static case. Thus, the approximate problem can be readily solved. Further, using arguments in Section 3.3, one can also derive a bound on the regret of this solution in a similar fashion to Proposition EC.2 in Online Appendix E.1. We skip the details for brevity.

Theorem 6 (Concavity of Dynamic Pricing Objective). *The objective function* (26a) *of the dynamic pricing Problem* (26) *is concave in* $\boldsymbol{\mu}$ *when* $0\le\beta\le 6.25/\left[\lambda_{\max}(\tilde{\mathbf{A}}+\tilde{\mathbf{A}}^\top)\cos\left(\frac{\pi}{T+1}\right)\right]$.

Finally, we conduct numerical experiments on the pricing problem. We begin by analyzing the static problem under two extreme scenarios: perfect price discrimination and uniform pricing. In both cases, we show that gradient-based algorithms yield near-optimal solutions. For the perfect price discrimination case, we extend our experiments to the dynamic problem and compare the results with static outcomes. Our findings reveal that a myopic pricing strategy is suboptimal. The optimal dynamic pricing strategy uses price to quickly guide consumer behavior toward the limiting behavior and then, stabilizes near the static optimum. Notably, during the transient period, the optimal price tends to increase, indicating that lower prices are used early to steer adoptions. A more detailed discussion of the implementation and results can be found in Online Appendices E.4 and E.5.

7. Conclusion

In this study, we focus on nonprogressive diffusion on a social network. We tide over the issues of the lack

of a general modeling framework and efficient algorithms in previous studies. Specifically, we base it on a general agent-based model incorporating local network effects and adaptable to various utility functions. We propose with a provable performance guarantee a fixed-point approximation scheme that can accurately and efficiently approximate the adoption behavior for all agents and validate the results through extensive experiments. We provide order-optimal bounds for the approximation error and conduct a thorough analysis of its dependency with network structure. Finally, we investigate the conventional optimization problems based on the fixed-point approximation.

We also view one of our contributions as proposing a novel approach to studying the user behavior within networks in a stochastic setting. In particular, there are several directions for future research, in which our method seems readily extendable. First, the adoptions may not change in each period but last for several periods in practice (e.g., a user needs to subscribe to Netflix for at least one month). It would be interesting to investigate the limiting behavior in this scenario. Second, this work only considers a binary-choice case, where each agent only decides to adopt or not. It is worth investigating whether similar results can be extended to a multiple-choice case (e.g., not to subscribe, to subscribe to a normal membership, or to subscribe to a premium membership). Finally, the local network effect is captured by the average adoption of the in-neighbors in our model. It is promising to consider the weighted average of in-neighbor adoptions where the network effect is asymmetric.

Acknowledgments

The authors acknowledge constructive feedback from Department Editor Jeannette Song, the anonymous associate editor, and three anonymous reviewers.

Endnotes

[1] Hereafter, we will refer to this as the nonprogressive LT model to differentiate it from its progressive counterpart.

[2] Although our primary results emphasize the limiting behavior, they can also be extended to capture the entire diffusion process, including transient behaviors.

[3] The CDF of logistic$(\mu;\sigma)$ is $F_\epsilon(x) = 1/(1+\exp\{-(x-\mu)/\sigma\})$.

[4] For some highly structured symmetric networks, solving the stationary distribution is tractable (see Online Appendix D.2).

[5] See https://networkrepository.com/networks.php.

[6] Negative prices are allowed, meaning that the platform can subsidize certain users, particularly those with a large influence on the network. In such cases, the platform incurs losses for these customers to generate a larger overall profit, which is a common practice.

References

Acemoğlu D, Como G, Fagnani F, Ozdaglar A (2013) Opinion fluctuations and disagreement in social networks. *Math. Oper. Res.* 38(1):1–27.

Acemoglu D, Dahleh MA, Lobel I, Ozdaglar A (2011) Bayesian learning in social networks. *Rev. Econom. Stud.* 78(4):1201–1236.

Afèche P, Liu Z, Maglaras C (2023) Ride-hailing networks with strategic drivers: The impact of platform control capabilities on performance. *Manufacturing Service Oper. Management* 25(5):1890–1908.

Agrawal S, Yin S, Zeevi A (2021) Dynamic pricing and learning under the bass model. *Proc. 22nd ACM Conf. Econom. Comput.* (ACM, New York), 2–3.

Allon G, Drakopoulos K, Manshadi V (2021) Information inundation on platforms and implications. *Oper. Res.* 69(6):1784–1804.

Anari N, Ehsani S, Ghodsi M, Haghpanah N, Immorlica N, Mahini H, Mirrokni VS (2013) Equilibrium pricing with positive externalities. *Theory Comput. Sci.* 476:1–15.

Ballester C, Calvó-Armengol A, Zenou Y (2006) Who's who in networks. Wanted: The key player. *Econometrica* 74(5):1403–1417.

Bapna R, Umyarov A (2015) Do your online friends make you pay? A randomized field experiment on peer influence in online social networks. *Management Sci.* 61(8):1902–1920.

Baron O, Hu M, Malekian A (2022) Technical note—Revenue volatility under uncertain network effects. *Oper. Res.* 70(4):2254–2263.

Bass FM (1969) A new product growth for model consumer durables. *Management Sci.* 15(5):215–227.

Benaïm M, Weibull JW (2003) Deterministic approximation of stochastic evolution in games. *Econometrica* 71(3):873–903.

Bonacich P (1987) Power and centrality: A family of measures. *Amer. J. Sociol.* 92(5):1170–1182.

Boykov Y, Veksler O, Zabih R (1998) Markov random fields with efficient approximations. *Proc. 1998 IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognition* (IEEE Computer Society, Washington, DC), 648–655.

Candogan O, Bimpikis K, Ozdaglar A (2012) Optimal pricing in networks with externalities. *Oper. Res.* 60(4):883–905.

Chandrasekhar AG, Larreguy H, Xandri JP (2020) Testing models of social learning on networks: Evidence from two experiments. *Econometrica* 88(1):1–32.

Chen N, Chen Y-J (2021) Duopoly competition with network effects in discrete choice models. *Oper. Res.* 69(2):545–559.

Chen Y, Shi C (2023) Network revenue management with online inverse batch gradient descent method. *Production Oper. Management* 32(7):2123–2137.

Chen X, van der Lans R, Trusov M (2021) Efficient estimation of network games of incomplete information: Application to large online social networks. *Management Sci.* 67(12):7575–7598.

Chen W, Wang Y, Yang S (2009) Efficient influence maximization in social networks. *Proc. 15th ACM SIGKDD Internat. Conf. Knowledge Discovery Data Mining* (ACM, New York), 199–208.

Chen W, Yuan Y, Zhang L (2010) Scalable influence maximization in social networks under the linear threshold model. *2010 IEEE Internat. Conf. Data Mining* (IEEE Computer Society, Washington, DC), 88–97.

Datareportal (2022) Digital 2022: Global overview report. Accessed June 27, 2022, https://datareportal.com/reports/digital-2022-global-overview-report.

Drakopoulos K, Zheng F (2017) Network effects in contagion processes: Identification and control. Columbia Business School Research Paper No. 18-8, Columbia University, New York.

Du C, Cooper WL, Wang Z (2016) Optimal pricing for a multinomial logit choice model with network effects. *Oper. Res.* 64(2):441–455.

Du C, Cooper WL, Wang Z (2018) Optimal worst-case pricing for a logit demand model with network effects. *Oper. Res. Lett.* 46(3):345–351.

Erdős P, Rényi A (1960) On the evolution of random graphs. *Math. Inst. Hungarian Acad. Sci.* 5(1):17–60.

Gallego G, Wang R (2014) Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities. *Oper. Res.* 62(2):450–461.

Göbel F, Jagers A (1974) Random walks on graphs. *Stochastic Processes Appl.* 2(4):311–336.

Goldenberg J, Libai B, Muller E (2001) Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Lett.* 12(3):211–223.

Golrezaei N, Jaillet P, Liang JCN (2020) No-regret learning in price competitions under consumer reference effects. *Annual Conf. Neural Inform. Processing Systems*, Advances in Neural Information Processing Systems, vol. 33 (Curran Associates Inc., Red Hook, NY), 21416–21427.

Gopalakrishnan M, Zhang H, Zhang Z (2023) Multiproduct pricing under the multinomial logit model with local network effects. *Decision Sci.* 54(4):447–466.

Granovetter M (1978) Threshold models of collective behavior. *Amer. J. Sociol.* 83(6):1420–1443.

Horst U, Scheinkman JA (2006) Equilibria in systems of social interactions. *J. Econom. Theory* 130(1):44–77.

Hu M, Wang Z, Feng Y (2020) Information disclosure and pricing policies for sales of network goods. *Oper. Res.* 68(4):1162–1177.

Huang J, Mani A, Wang Z (2022) The value of price discrimination in large social networks. *Management Sci.* 68(6):4454–4477.

Ising E (1924) Beitrag zur theorie des ferro-und paramagnetismus. PhD thesis, Grefe & Tiedemann, Hamburg, Germany.

Jackson MO (2010) *Social and Economic Networks* (Princeton University Press, Princeton, NJ).

Jackson MO, Lin Z, Yu NN (2020) Adjusting for peer-influence in propensity scoring when estimating treatment effects. Preprint, submitted January 20, http://dx.doi.org/10.2139/ssrn.3522256.

Jadbabaie A, Molavi P, Sandroni A, Tahbaz-Salehi A (2012) Non-Bayesian social learning. *Games Econom. Behav.* 76(1):210–225.

Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. *Proc. Roy. Soc. London A* 115(772): 700–721.

Kempe D, Kleinberg J, Tardos É (2003) Maximizing the spread of influence through a social network. Getoor L, Senator TE, Domingos PM, Faloutsos C, eds. *Proc. Ninth ACM SIGKDD Internat. Conf. Knowledge Discovery Data Mining* (ACM, New York), 137–146.

Li H (2020) Optimal pricing under diffusion-choice models. *Oper. Res.* 68(1):115–133.

Li H, Huh WT (2011) Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications. *Manufacturing Service Oper. Management* 13(4):549–563.

Li Y, Fan J, Wang Y, Tan K-L (2018) Influence maximization on social graphs: A survey. *IEEE Trans. Knowledge Data Engrg.* 30(10):1852–1872.

Lin Y, Wang M, Zhang H, Zhang R, Shen Z-JM (2024) Content promotion for online content platforms with the diffusion effect. *Manufacturing Service Oper. Management* 26(3):1062–1081.

Lu Y, Jerath K, Singh PV (2013) The emergence of opinion leaders in a networked online community: A dyadic model with time dynamics and a heuristic for fast estimation. *Management Sci.* 59(8):1783–1799.

Ma L, Krishnan R, Montgomery AL (2015) Latent homophily or social influence? An empirical analysis of purchase within a social network. *Management Sci.* 61(2):454–473.

Mislove AE (2009) Online social networks: Measurement, analysis, and applications to distributed information systems. PhD thesis, Rice University, Houston, TX.

Nemhauser GL, Wolsey LA, Fisher ML (1978) An analysis of approximations for maximizing submodular set functions—I. *Math. Programming* 14(1):265–294.

Nosrat F, Cooper WL, Wang Z (2021) Pricing for a product with network effects and mixed logit demand. *Naval Res. Logist.* 68(2):159–182.

Rheinboldt WC (1998) *Methods for Solving Systems of Nonlinear Equations* (SIAM, Philadelphia).

Rossi RA, Ahmed NK (2015) The network data repository with interactive graph analytics and visualization. *Proc. Twenty-Ninth AAAI Conf. Artificial Intelligence* (AAAI Press, Palo Alto, CA).

Sadler E (2020) Diffusion games. *Amer. Econom. Rev.* 110(1):225–270.

Schelling TC (1978) *Micromotives and Macrobehavior* (Norton, New York).

Shriver SK, Nair HS, Hofstetter R (2013) Social ties and user-generated content: Evidence from an online social network. *Management Sci.* 59(6):1425–1443.

Song J-S, Zipkin P (2009) Inventories with multiple supply sources and networks of queues with overflow bypasses. *Management Sci.* 55(3):362–372.

Van Mieghem P, Omic J, Kooij R (2009) Virus spread in networks. *IEEE/ACM Trans. Networking* 17(1):1–14.

Wang R, Wang Z (2017) Consumer choice models with endogenous network effects. *Management Sci.* 63(11):3944–3960.

Xie T, Wang Z (2025) Personalized pricing and assortment optimization under consumer choice models with local network effects. *Oper. Res.* 73(3):1289–1306.

Yang N, Zhang RP (2022) Dynamic pricing and inventory management in the presence of online reviews. *Production Oper. Management* 31(8):3180–3197.

Yeomans JM (1992) *Statistical Mechanics of Phase Transitions* (Oxford Academic, Oxford, UK).